



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2020

Méthodes et outils numériques pour l'analyse avancée des couleurs des films

Flückiger, Barbara ; Halter, Gaudenz

Abstract: Les couleurs sont l'élément du style filmique qui posent le plus grand défi, mais elles sont aussi l'élément le plus gratifiant à investiguer. Longtemps mises à l'écart des études de cinéma, les couleurs des films connaissent un regain d'intérêt depuis une dizaine d'années. S'appuyant sur les avancées récentes en matière d'analyse à l'aide de bases de données, d'outils d'apprentissage profond et de méthodes de datavisualisation, le projet de recherche ERC Advanced Grant FilmColors a comme objectif l'élaboration d'une approche globale des multiples aspects de la couleur des films. Ce chapitre met l'accent sur un ensemble de concepts théoriques et analytiques des couleurs des films –y compris l'interprétation humaine–, reliant les dimensions stylistiques, expressives et narratives au développement et à l'évaluation de méthodes numériques. Depuis 2017, un corpus de plus de 400 films a été analysé avec un procédé computationnel mis en œuvre dans le logiciel d'annotation et d'analyse filmique VIAN. Ce logiciel est par ailleurs connecté à la plateforme en ligne VIAN Webapp dans le but de faciliter les requêtes, la visualisations et l'évaluation de résultats à plusieurs niveaux : séquences, films et corpus. Comparé aux approches traditionnelles de l'esthétique, de la technologie et de la narratologie des couleurs des films (largement basées sur une tradition linguistique), les outils des humanités numériques nous permettent de représenter instantanément les résultats à travers une grande variété de modèles visuels. En outre, ces représentations visuelles peuvent être augmentées d'annotations et d'interprétations apportées par les utilisateurs eux-mêmes. Notre approche essaie de consolider un écosystème global dédié à la recherche sur l'esthétique et la narrativité filmique, en contribuant aux méthodes existantes des études de cinéma.

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-199220>

Book Section

Published Version



The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:

Flückiger, Barbara; Halter, Gaudenz (2020). Méthodes et outils numériques pour l'analyse avancée des couleurs des films. In: Le Cor, Gwen; Reyes, Everardo. Faire image : inscriptions, espaces et formes numériques = What makes an image? : inscriptions, digital spaces and forms. Saint-Denis: Octaviana, 63-129.

Faire image

Inscriptions,
espaces et
formes numériques

SOUS LA DIRECTION DE
Gwen Le Cor
et **Everardo Reyes**

Ce travail a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme d'Investissements d'avenir portant la référence ANR-17-EU- RE-0008
www.eur-artec.fr



ISBN 978 23 7059 012 1

Image de couverture : carte thermique des couleurs produit avec VIAN WebApp, Barbara FLUECKIGER et Gaudenz HALTER

Design et mise en page : Everardo REYES

Police : CMU Serif (OFL, <https://fontlibrary.org/en/font/cmu-serif>)



Creative Commons License CC BY-NC-ND <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Attribution - Pas d'Utilisation Commerciale - Pas de Modification

© Les auteurs / Octaviana - Bibliothèque numérique Paris 8, Saint-Denis, décembre 2020

Vous êtes autorisé à :

Partager — copier, distribuer et communiquer le matériel par tous moyens et sous tous formats

Selon les conditions suivantes :

Attribution — Vous devez créditer l'Œuvre, intégrer un lien vers la licence et indiquer si des modifications ont été effectuées à l'Œuvre. Vous devez indiquer ces informations par tous les moyens raisonnables, sans toutefois suggérer que l'Offrant vous soutient ou soutient la façon dont vous avez utilisé son Œuvre.

Pas d'Utilisation Commerciale — Vous n'êtes pas autorisé à faire un usage commercial de cette Œuvre, tout ou partie du matériel la composant.

Pas de modifications — Dans le cas où vous effectuez un remix, que vous transformez, ou créez à partir du matériel composant l'Œuvre originale, vous n'êtes pas autorisé à distribuer ou mettre à disposition l'Œuvre modifiée.

Notes:

Vous n'êtes pas dans l'obligation de respecter la licence pour les éléments ou matériel appartenant au domaine public ou dans le cas où l'utilisation que vous souhaitez faire est couverte par une exception.

Faire Image / What makes an image?

*Inscriptions, espaces et formes numériques /
Inscriptions, digital spaces and forms*

Sous la direction / Edited by
Gwen Le Cor & Everardo Reyes

Octaviana - Bibliothèque numérique Paris 8
Saint-Denis

Table de matières / Contents

	Introduction	
	<i>Gwen Le Cor & Everardo Reyes</i>	1
1	Drawing the Line: On the Making of the Images in <i>The Pleasure of the Coast: A Hydro-graphic Novel</i>	
	<i>J.R. Carpenter</i>	13
2	Interfaces et flâneries numériques : poétique de l'œil distrait dans les desktops documentaires de Kevin B. Lee	
	<i>Marie-Pierre Burquier</i>	29
3	Regard sur les nouvelles œuvres narratives numériques - art et narrativité	
	<i>Véronique Peyrot</i>	47
4	Méthodes et outils numériques pour l'analyse avancée des couleurs des films	
	<i>Barbara Flueckiger & Gaudenz Halter</i>	63
5	Inconceivable Yet Computable: Algorithmic and Sound Art Based on the High Tide at the Venice Biennale	
	<i>Francesca Franco</i>	131
6	Bref retour sur le parcours des images ; des habitudes visuelles à l'exploration politique, les perspectives numériques	
	<i>Manon Denoun</i>	139
7	Images circonstanciées	
	<i>Florent Di Bartolo</i>	157
8	Topology of augmented spaces	
	<i>Federico Biggio</i>	171
9	<i>Flânerie 2.0</i> – Vidéo numérique, 11 minutes, 2018	
	<i>Chloé Galibert-Lainé</i>	183

4

Méthodes et outils numériques pour l'analyse avancée des couleurs des films¹⁷

Barbara Flueckiger et Gaudenz Halter

Abstract

Les couleurs sont l'élément du style filmique qui posent le plus grand défi, mais elles sont aussi l'élément le plus gratifiant à investiguer. Longtemps mises à l'écart des études de cinéma, les couleurs des films connaissent un regain d'intérêt depuis une dizaine d'années. S'appuyant sur les avancées récentes en matière d'analyse à l'aide de bases de données, d'outils d'apprentissage profond et de méthodes de datavisualisation, le projet de recherche ERC Advanced Grant *FilmColors* a comme objectif l'élaboration d'une approche globale des multiples aspects de la couleur des films. Ce chapitre met l'accent sur un ensemble de concepts théoriques et analytiques des couleurs des films –y compris l'interprétation humaine–, reliant les dimensions stylistiques, expressives et narratives au développement et à l'évaluation de méthodes numériques.

Depuis 2017, un corpus de plus de 400 films a été analysé avec un procédé computationnel mis en œuvre dans le logiciel d'annotation et d'analyse filmique VIAN. Ce logiciel est par ailleurs connecté à la plateforme en ligne VIAN Webapp dans le but de faciliter les requêtes, la visualisations et l'évaluation de résultats à plusieurs niveaux : séquences, films et corpus. Comparé aux approches traditionnelles de l'esthétique, de la technologie et de la narratologie des couleurs des films (largement basées sur une tradition linguistique), les outils des humanités numériques nous permettent de représenter instantanément les résultats à travers une grande variété de modèles visuels. En outre, ces représentations visuelles peuvent être augmentées d'annotations et d'interprétations apportées par les utilisateurs eux-mêmes. Notre approche essaie de consolider un écosystème global dédié à la recherche sur l'esthétique et la narrativité filmique, en contribuant aux méthodes existantes des études de cinéma.

¹⁷ Titre original : Flueckiger, Barbara & Halter, Gaudenz (2020) "Methods and Advanced Tools for the Analysis of Film Colors in Digital Humanities", In *Digital Humanities Quarterly*, 14,4, Special Issue Digital Humanities & Film Studies. Analyzing the Modalities of Moving Images. Traduit de l'anglais par Achilleas Papakonstantis.

Mots clés : couleurs des films, analyse filmique, outils numériques, représentation visuelle.

1 Introduction

La couleur constitue l'un des éléments stylistiques des films parmi les plus difficiles à analyser, mais également – comme nous essaierons de le montrer dans ce chapitre – parmi les plus enrichissants. Longtemps négligées dans le champ des études cinématographiques, les couleurs des films ont fait l'objet d'une attention croissante au cours de la dernière décennie. Suite à l'avènement récent des méthodes numériques, des études approfondies sur leur histoire, leurs utilisations, leurs concepts sous-jacents, ainsi que des réflexions théoriques et épistémologiques en humanités numériques ont été publiées par Olivia Kristina Stutz (2016), Adelheid Heftberger (2016, 2018) et Christian Gosvig Olesen (2017). On assiste également de plus en plus à une circulation des discours autour du développement d'approches, de méthodes et d'outils numériques pour l'analyse de film. Ainsi, Stutz (2016) et Olesen (2017) portent une attention particulière à la question de l'analyse de la couleur. Kevin Ferguson (2013, 2015), Lev Manovich (2013, 2015) et Everardo Reyes (2014, 2017) ont développé des méthodes de visualisation spécifiques pour les œuvres d'art, les tableaux et les films. Enfin, la publication de Pause et Walkowski sur l'analyse des couleurs assistée par ordinateur s'inspire de notre propre travail (Pause/Walkowski, 2018).

Les analyses traditionnelles des couleurs des films étaient principalement basées sur une description verbale. Elles affichaient une tendance à l'interprétation herméneutique, tandis que les dimensions esthétiques et affectives restaient souvent négligées.

Avec le développement d'analyses à partir d'une base de données, d'outils d'apprentissage profond et d'une large gamme de méthodes de visualisation, le projet de recherche ERC Advanced Grant *FilmColors* (voir les *Remerciements* dans la Section 11) s'attache à fournir une approche plus complète de l'analyse des couleurs des films sous leurs multiples aspects.

Ainsi, l'argument central de ce chapitre met l'accent sur un ensemble de concepts théoriques et analytiques solides portant sur les couleurs des films – y compris l'interprétation humaine – qui relie les dimensions stylistique, expressive et narrative au développement et à l'évaluation des méthodes numériques.

Selon une idée fausse mais néanmoins répandue, les outils numériques génèrent des résultats significatifs de manière automatique. Or, un raisonnement théorique solide et une réflexion constante sur les méthodes de visualisation et leurs fondements épistémologiques et perceptuels constituent des exigences obligatoires régissant toute collaboration interdisciplinaire entre le champ des études cinématographiques et la science informatique – pour une discussion plus approfondie sur ces conditions préalables et leur lien avec l'esthétique expérimentale, voir nos recherches précédentes (Flueckiger 2011, 2017 ; Flueckiger/Halter 2018). Contrairement à ces dernières, la présente publication vise à apporter un aperçu de la variété des méthodes employées, des obstacles, des avancées et des problèmes rencontrés, ainsi que des leçons tirées durant ce déploiement collaboratif des outils.

Une partie intégrante de nos projets de recherche est la *Timeline of Historical Film Colors* (<https://filmcolors.org/>), une ressource Web interactive et abondamment documentée sur les couleurs des films, créée et gérée par Barbara Flueckiger depuis 2012.

2 Analyse à partir d'une base de données, concepts analytiques et évaluation

L'objectif principal de la perspective interdisciplinaire de notre projet de recherche est d'étudier la relation entre l'esthétique et la technologie des couleurs des films. À cette fin, un groupe important de plus de quatre cents films produits entre 1895 et 1995 a été analysé de manière extrêmement détaillée via une approche assistée par ordinateur. Cette approche combinait la segmentation temporelle à l'aide de l'outil d'annotation vidéo ELAN (dont la première version a été commercialisée en 2002) avec un protocole géré par une base de données, consistant en un vocabulaire contrôlé d'environ 1200 concepts théoriques et analytiques, employé pour l'annotation de chaque segment. Un réseau de bases de

données relationnelles (figure 1), constituées sur mesure pour l'analyse, les données filmographiques, le glossaire et l'évaluation des résultats a été programmé sur FileMaker, en grande partie par la chercheuse principale (CP) et avec la contribution de son équipe. Malgré les faiblesses et les limites de FileMaker en termes de programmabilité et de flexibilité, l'adaptation et le développement autosuffisants des bases de données en réponse aux exigences évolutives des analyses sont restés un atout important tout au long du projet. De plus en plus, les bases de données étaient liées les unes aux autres par le biais de connexions relationnelles afin de produire des résultats significatifs et de fournir un accès direct à une variété de méthodes d'évaluation des données récoltées (voir Section 3).

Notre équipe a distingué plusieurs niveaux d'analyse : des captures d'écran aux segments temporels de films individuels (*niveau micro*), des films entiers (*niveau méso*) et le corpus dans sa totalité ou des sous-corpus sélectionnés (*niveau macro*) – voir Olesen (2016).

Nous avons collecté des données filmographiques afin de représenter l'ensemble du corpus, de déterminer sa sélection ainsi que son attribution à des chercheurs individuels, et d'assurer le suivi du statut de traitement. Les corpus ont été sélectionnés à partir de recherches sur des monographies et des articles traitant des couleurs des films, chacune des doctorantes définissant son centre d'intérêt parmi les trois périodes retenues : 1895-1930, 1930-1955 et 1955-1995. La ligne directrice suivie pour cette sélection vise à comparer des œuvres canoniques, célébrées pour leur utilisation élaborée ou audacieuse des couleurs, l'idée étant de former des sous-catégories selon les procédés spécifiques de couleur, les genres, les cinéastes, les chefs opérateurs, les consultants en couleur ou les pays. Conformément à la poétique historique (Bordwell 1989) et à l'analyse néo-formaliste établie par les représentants de l'école de Wisconsin, Kristin Thompson et David Bordwell (Thompson 1988 ; Bordwell et al. 1985), les corpus sélectionnés devraient permettre l'*analyse diachronique* des styles personnels, des contextes institutionnels – par exemple, l'évolution des normes professionnelles, des préférences culturelles, de la notion de goût – ou des innovations technologiques au fil du temps, ainsi qu'une

comparaison synchronique parmi ces différentes instances au cours d'une période donnée.

L'identification du type de pellicule, la recherche sur l'innovation technique et les informations détaillées sur les procédés de couleur appliqués à chaque film jouent un rôle important dans la promotion d'une meilleure compréhension de ces connexions et permettent de contourner certaines idées fausses présentes dans la littérature antérieure. Ces méthodes de recherche sont complétées par l'évaluation scientifique des caractéristiques spectrales des types de pellicule, afin de mieux faciliter la numérisation et la restauration des couleurs des films analogiques. Une telle approche englobante reliant les connaissances sur les évolutions esthétiques à une étude approfondie de l'innovation technologique a été nommée « technobole » par Frank Beau (2002). Contrairement au déterminisme technique, l'approche « technobole » qui se trouve au centre de notre méthode prend en considération les facteurs épistémologiques, institutionnels, sociaux, culturels et économiques qui régissent les avancées techniques et la manière dont la technologie à son tour se répercute sur la culture et la société. Cette perspective culturelle est étudiée en profondeur dans le deuxième projet de recherche de la chercheuse principale (CP), *Film Colors. Technologies, Cultures, Institutions*.

Un modèle à trois niveaux conforme aux schémas de métadonnées recommandés, établis pour les archives cinématographiques par des entités de normalisation telles la FIAF ou filmstandards.org, a été mis en œuvre dans la base de données du corpus par une des membres de notre équipe, Joëlle Kost. Il permet l'attribution de fiches individuelles pour un seul film, telles que les DVD, les Blu-ray ou divers tirages et négatifs inspectés et documentés dans les archives cinématographiques en Europe, au Japon et aux Etats-Unis, et attribue un identifiant d'article (*Item ID*) tripartite, spécifique à chacune d'elles. Dès le départ, cette base de données était hébergée sur un serveur FileMaker, mis à disposition par l'Université de Zurich.

Tous les films choisis pour l'analyse ont été numérisés, puis segmentés temporellement à l'aide de l'outil d'annotation ELAN (pour une description de notre approche, voir la Section 4, *Développement d'une*

plateforme d'annotation visuelle, d'analyse et de visualisation). Les informations obtenues – principalement des *time codes* et des segments numérotés ainsi que des descriptions élémentaires facultatives à l'aide d'un modèle – ont ensuite été exportées vers une base de données d'analyse (*Analysis DB*) pour une lecture attentive, avant d'être annexées à la base de données de corpus (*Corpus DB*) pour les données filmographiques, en fonction de l'ID de l'élément du film en considération.

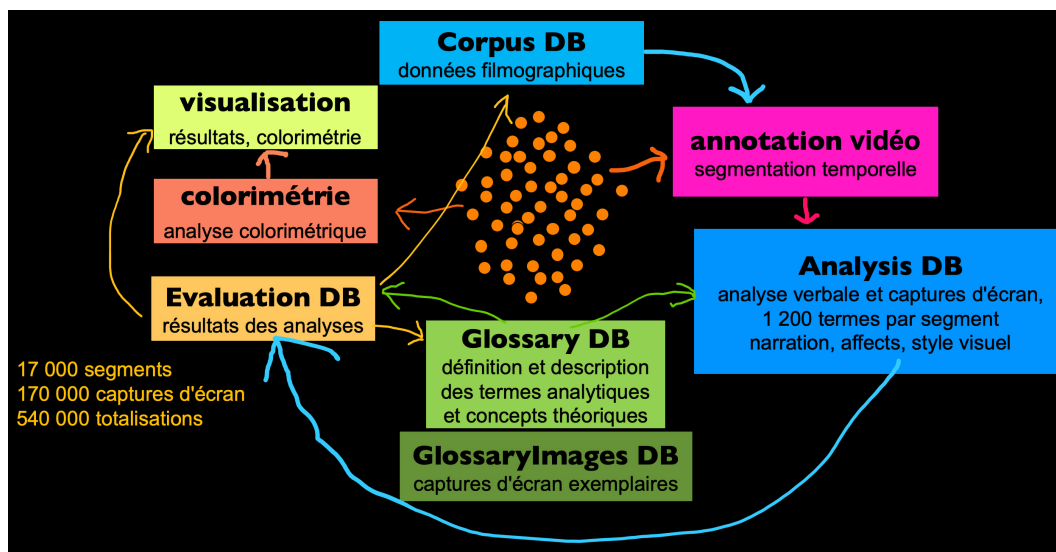


Figure 1. Schéma graphique du flux opérationnel d'analyse et d'évaluation avec l'architecture de bases de données

Les concepts théoriques et analytiques ont en grande partie été élaborés avant le début du projet, au cours des activités d'enseignement et de recherche de la CP dans les domaines de la théorie des couleurs, de l'esthétique de film, de la sémiologie et de la narratologie – l'accent étant mis sur les couleurs des films au cours de la dernière décennie. Par conséquent, ils faisaient déjà partie de la proposition de recherche soumise à l'ERC (Conseil européen de la recherche). Pendant une période de six mois de coaching et de formation au début du projet, les membres de l'équipe ont été initiés aux concepts et ont eu l'occasion de participer à de longues discussions. Certains concepts ont émergé au cours des analyses suivant une approche ascendante (*bottom-up*), apportés par les membres de l'équipe sur la base de leurs observations. Par exemple, un catalogue des termes élémentaires pour les états affectifs ou émotionnels des personnages faisait déjà partie du glossaire initial mais s'est trouvé sans cesse élargi par la chercheuse postdoctorale Bregt Lameris qui se spécialise

sur la relation entre les couleurs des films et les affects. Des motifs et des thèmes ont également émergé selon une approche ascendante pour certaines situations canoniques – par exemple, les cérémonies, les numéros de spectacle, les métamorphoses – comme d’ailleurs des thématiques tels l’exotisme, l’architecture, l’auto-réflexivité, etc. Ils ont été organisés dans une base de données de mots-clés, annexée à la base de données d’analyse.

Onze registres différents comprenaient une taxonomie des classes de concepts analytiques, allant des descriptions verbales des couleurs, des contrastes de couleurs, de la composition de l’image, de la profondeur de champ, de l’éclairage, des textures et des motifs, des propriétés de surface et des matériaux des personnages – y compris leurs propriétés tactiles – à la matérialité des films analysés suivant le concept de *faktura* (Flueckiger 2020), en passant par les mouvements de caméra, des personnages, des objets et des lumières.

Pour chacun des segments temporels – il y avait généralement entre cinquante et soixante-dix segments par film –, l’équipe a parcouru toute la gamme de concepts analytiques dans la base de données d’analyse et a mis en ligne jusqu’à trente-deux captures d’écran pertinentes en formats compressés.

Tous les concepts théoriques et analytiques du vocabulaire contrôlé – plus de 1200 termes analytiques, y compris les teintes – étaient continuellement définis dans une base de données de glossaire (*Glossary DB*) avec des références aux sources, quand elles étaient disponibles, et illustrés avec des captures d’écran représentatives rassemblées lors des analyses.

Dès le départ, nous avons réfléchi à la manière dont ces termes analytiques pouvaient être traités avec des outils avancés d’extraction automatique de données, et cette préoccupation continue à guider le développement d’outils d’apprentissage profond (voir Sections 4 à 9). Il est évident que tous les problèmes ne peuvent pas être facilement résolus avec la gamme limitée des ressources de pointe actuelles, même dans le cadre d’un projet de recherche ERC de cette envergure. D’autre part, comme indiqué ci-dessus, c’est le credo central de l’approche englobante du projet de viser à une analyse qualitative centrée sur la contextualisation des

observations par l'intervention et l'interprétation humaines. De par leur nature même, les approches automatiques du traitement d'image ne sont pas en mesure d'identifier des détails subtils et des idiosyncrasies – par exemple, le fait que les rideaux bougent légèrement avec le vent pourrait être une métaphore de l'agitation intérieure de l'héroïne, comme dans le film japonais JIGOKUMON.

Les **concepts narratologiques** posent un défi particulier à l'évaluation automatisée. Les structures de point de vue qui fonctionnent avec le concept de *focalisation* (Genette 1972, 1983) comme moyen de différenciation parmi les instances de narration – les soi-disant focalisateurs ou filtres – peuvent s'avérer très difficiles à identifier pour des observateurs non-humains, mais elles sont extrêmement importantes pour l'étude des couleurs des films, y compris les états mentaux des personnages dans les rêves ou les hallucinations, l'alignement sur les personnages (Smith 1995), l'organisation temporelle de la narration comme les flashbacks, les sommaires, la mise en abyme, le montage parallèle ou les séquences de montage, les situations non-narratives, les points tournants, le suspense et la préfiguration, etc.

Des défis similaires se posent pour les phénomènes de **sémantique d'ordre supérieur** – en d'autres termes, toutes les formes de modification du sens qui sont établies via des relations intra-textuelles, des références intertextuelles et inter-médiatiques, ou des emplois culturels, tels les contextes culturels, le milieu, le goût, les marqueurs sociopolitiques, les stéréotypes, les conventions de genre, les relations entre les personnages, le genre (*gender*), les symboles, les signaux, les métaphores et les isotopies/les rimes – les références intertextuelles et inter-médiatiques se liant à d'autres films, médias ou œuvres d'art via le pastiche, l'allusion, la citation, l'ironie, la parodie, etc. (Genette 1992 ; Jameson 1991 ; Dyer 2006).



Figure 2. Des sémantiques d'ordre supérieur dans le traitement de la couleur : la pratique du pastiche dans *BLOOD AND SAND* [ARÈNES SANGLANTES] (États-Unis 1941, Rouben Mamoulian) (en haut) et des marqueurs sociopolitiques indiquant le statut et la période dans *THE PRIVATE LIVES OF ELIZABETH AND ESSEX* [LA VIE PRIVÉE D'ELISABETH D'ANGLETERRE] (États-Unis 1939, Michael Curtiz) (en bas).

Les émotions et les affects se rapportent non seulement aux états intérieurs des personnages – joie, tristesse, colère, haine, dégoût – ou aux relations émotionnelles centrées sur l’amour, les conflits, le sexe, etc., mais également à des relations transmodales véhiculées par les représentations visuelles de l’odorat, du goût ou du toucher. En outre, il existe des concepts théoriques de base pour les réponses émotionnelles et affectives telles que l’affect direct (Plantinga 2009), la contagion et le mimétisme, ainsi que des catégories esthétiques qui interpellent les sens tels que l’excès (Thompson 1986) (figure 3), l’émotion d’artefact (Tan 1996), l’ambiance ou l’humeur (*Stimmung*).

Or, dans le domaine de **l’identification des couleurs**, les approches informatiques sont supérieures à l’observation humaine, même s’il est impératif de souligner le fait que les résultats de ces analyses nécessitent également une interprétation humaine.

Dans notre base de données d’analyse, les couleurs ont été décrites verbalement en tant que **teintes dominantes** pour la scène entière, les protagonistes féminins et masculins et les personnages secondaires, l’arrière et le premier plan, les intertitres et les lettres, en y incluant des observations générales sur la saturation, la clarté, etc. Il est évident que de telles descriptions verbales sont très limitées car elles ne prennent pas en compte les nuances subtiles de chaque couleur à l’intérieur d’une certaine gamme, exprimées en différents types de teintes avec des niveaux différents de saturation ou de luminosité.

Par **schémas de couleurs** – souvent appelés *palettes de couleurs* – nous désignons la distribution globale des couleurs dans une image ou dans un segment temporel en fonction des variations de teintes, de saturation, de chaleur ou de clarté, telles que monochrome, restrictive, tamisée, criarde, saturée, etc. Ces types peuvent se produire simultanément : par exemple, un schéma monochrome peut être en même temps chaud ou saturé. Comme nous le développerons dans une section ultérieure (voir Section 7, *Analyses colorimétriques et visualisations*), nos méthodes d’identification des schémas de couleurs à l’aide d’outils d’apprentissage profond (*deep learning*) sont à la fois supérieures et plus raffinées que les

descriptions verbales et procurent des résultats fortement significatifs, tant qu'ils sont, encore une fois, reliés aux concepts développés ci-dessus.

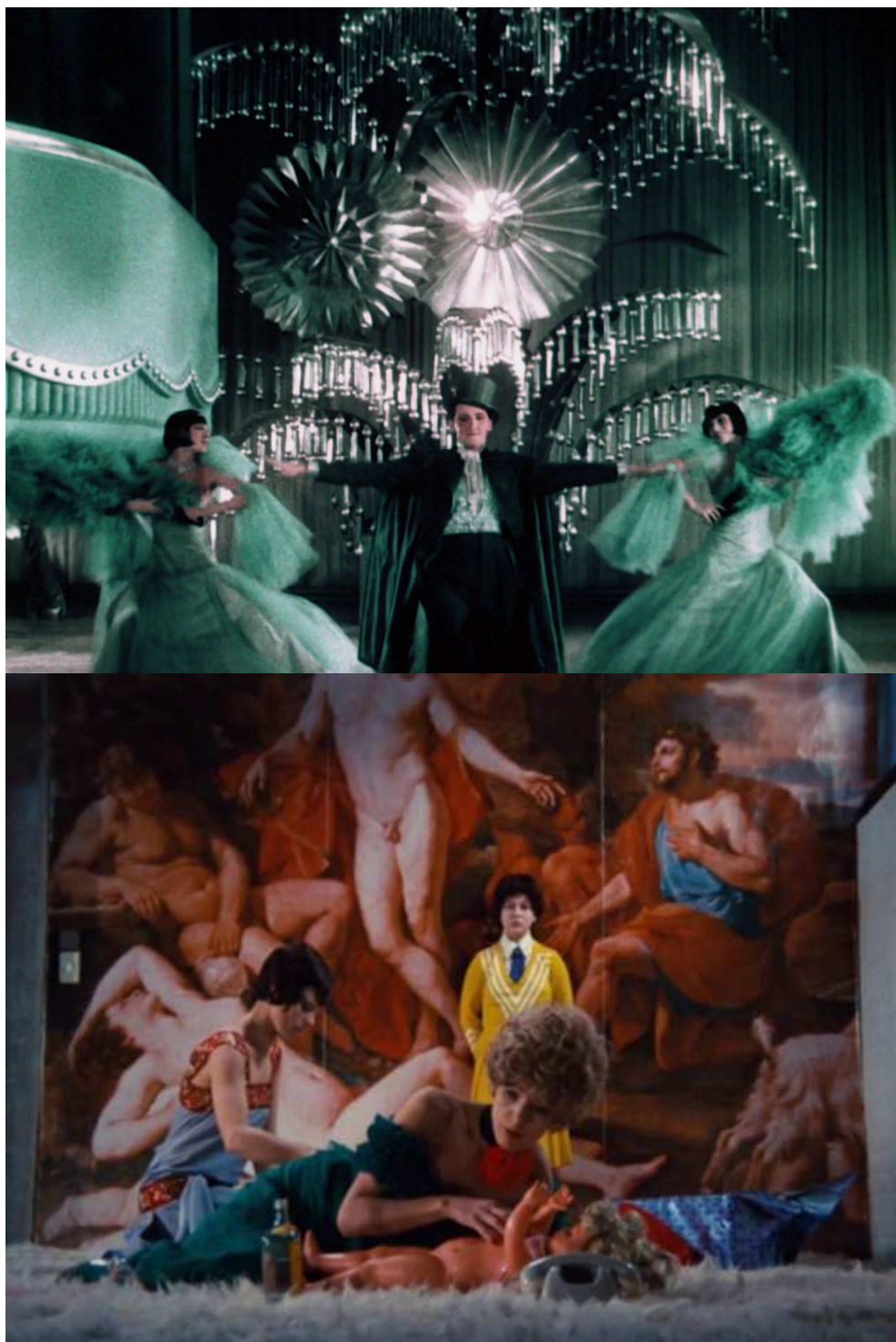


Figure 3. L'excès dans KING OF JAZZ [LA FÉERIE DU JAZZ] (États-Unis 1930, John Murray Anderson; Pál Fejös) (en haut) et DIE BITTEREN TRÄNEN DER PETRA VON KANT [LES LARMES AMÈRES DE PETRA VON KANT] (Allemagne 1972, Rainer Werner Fassbinder) (en bas).

Les **contrastes de couleurs** renvoient à un ensemble déjà établi, déterminé par l'artiste et universitaire Johannes Itten (1970) dans le but de décrire les relations spécifiques des harmonies de couleurs, de leurs « accords » ou de leur distribution spatiale, corrélées de nouveau aux dimensions de la teinte, de la saturation et de la clarté telles qu'organisées par Itten dans *Farbstern (L'Étoile des couleurs)*. Par exemple, le contraste de couleur le plus omniprésent au cours de la dernière décennie a été la combinaison orange-bleu sarcelle (*orange-teal*) qui est à la fois un contraste chaud-froid et un contraste complémentaire, tout en contenant une variation clair-obscur, puisque le jaune est perçu comme plus clair que le bleu.

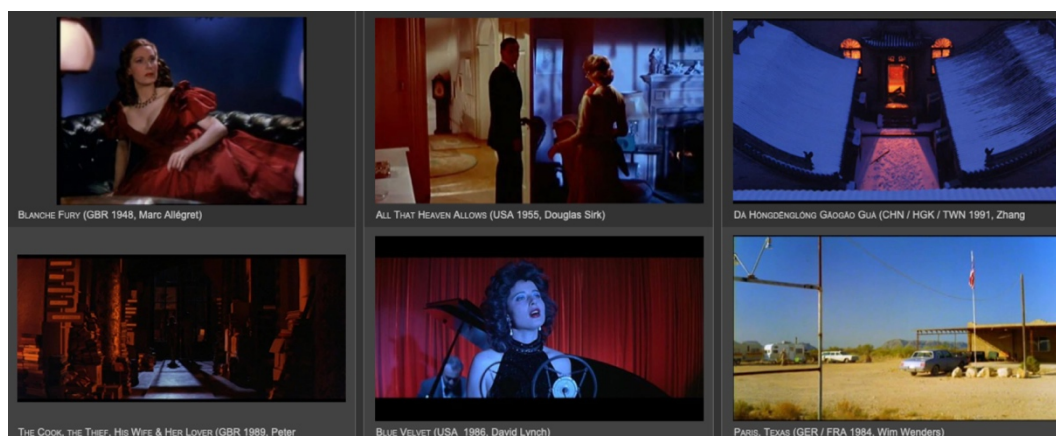


Figure 4. Une sélection de captures d'écran représentatives du concept « contraste chaud-froid » dans la base de données de glossaire.

L'identification des contrastes de couleurs et des schémas de couleurs est depuis longtemps un des domaines de l'analyse informatique. Cependant, comme nous le verrons plus loin (voir Section 7, *Analyses colorimétriques et visualisations*), la plupart de ces approches ne sont pas conformes aux exigences de l'analyse esthétique en termes de différenciation, de flexibilité et de subtilité. Certaines d'entre elles ont été établies à des fins normatives, afin d'identifier les « bonnes » utilisations des harmonies de couleurs pour le design ou pour donner une visualisation approximative, quoique agréable, destinée aux cinéphiles, comme dans MovieBarcodes.

2.1 Problèmes

Le plus grand défi a toujours été le niveau de complexité pour tous les membres de l'équipe travaillant sur les analyses des films. Dans l'ensemble,

le processus a été perçu comme extrêmement chronophage. Après une première évaluation effectuée plusieurs mois après le lancement du projet, les concepts ont été séparés en deux catégories : ceux qui étaient les plus pertinents et ceux qui restaient plus rares. Les membres de l'équipe ont également rencontré des difficultés en travaillant sur des catalogues de concepts si vastes qu'ils n'étaient organisés en fonction de leur pertinence que de manière aléatoire. Par conséquent, ils ont conçu des listes ordonnées triées selon la cohérence thématique. Enfin, nous avons fini par établir des dispositions graphiques individuelles pour chaque membre de l'équipe afin de tenir compte de leurs centres d'intérêt respectifs. Néanmoins, cette approche a donné lieu à une architecture bien plus complexe quant à la collecte et l'évaluation de toutes les données.

2.2 Leçons tirées

Nourrie par une architecture de flux opérationnels et de bases de données en constante évolution, la plateforme de *crowdsourcing* pour les utilisateurs externes (Flueckiger/Halter 2018 ; Halter et al. 2019) a été développée de manière modulaire, à nouveau par Gaudenz Halter avec l'aide de Silas Weber. Dans un premier temps, les concepts sont triés en fonction des relations internes – par exemple, les affects positifs liés à la joie *vs* les affects négatifs liés à la dépression ou l'agression, tels qu'ils ont été élaborés et affinés pendant le développement du flux opérationnel et à la suite de l'évaluation finale. Dans un deuxième temps, des niveaux d'importance et de complexité conceptuelles ont été établis pour chaque domaine de l'analyse, tels l'éclairage ou la composition de l'image. Ce design modulaire donnera aux utilisateurs la possibilité de sélectionner dans le menu de concepts non seulement les thématiques qui les intéressent – par exemple, les contrastes de couleurs ou les costumes – mais également le niveau de complexité de chacun de ces modules examiné séparément, de sorte que le vocabulaire contrôlé corresponde le mieux aux centres d'intérêt de leur recherche.

3 Évaluation des données récoltées

Le processus d'analyse manuelle a donné lieu à une énorme quantité de données et de captures d'écran : plus de 17000 segments avec environ

170000 captures d'écran assemblés dans une base de données d'analyse principale (*Master Analysis DB*) et plus d'un demi-million de compilations dans une base de données d'évaluation. Cette **base de données d'évaluation** a été reliée à la **base de données de glossaire** et à la **base de données de corpus**, sur la base de l'ID de glossaire et l'ID de l'élément du film (figure 1). Avec ces identifiants, nous avons pu ensuite afficher les résultats directement dans la base de données de glossaire et la base de données de corpus respectivement en utilisant des portails et des scripts, ce qui s'est avéré être le plus grand avantage de l'architecture relationnelle des bases de données. Comme nous le verrons ci-dessous (voir Section 3.1), il y avait également de nombreux problèmes et défis à maîtriser.

En principe, les résultats sont accessibles de trois manières via l'architecture FileMaker et de nombreuses autres manières, plus complexes, via l'outil d'analyse et d'annotation VIAN et la plateforme en ligne VIAN WebApp. VIAN WebApp est un portail de *crowdsourcing* contenant actuellement tous les films analysés à ce jour (plus de quatre cents) : ici, ils peuvent être évalués et visualisés au niveau du segment, du film et du corpus (Flueckiger/Halter 2018 ; Halter et al. 2019). Dans l'avenir, les utilisateurs externes auront la possibilité de mettre en ligne leurs propres projets VIAN. Étant donné que la base de données de VIAN WebApp héberge à la fois des informations qualitatives et numériques sur les couleurs, les développeurs ont opté pour une architecture diploïde de bases de données pour la plateforme en ligne. La plupart des données sont hébergées sur une base de données Postgres SQL ; pour des recherches rapides et le traitement d'informations numériques, ils utilisent une structure de fichiers HDF5 (voir Halter et al. 2019). Les données sont traitées par *cloud computing* sur Microsoft Azure.

L'analyse VIAN hors ligne est intégrée dans un écosystème (voir figure 5). Les projets individuels sont mis sur la plateforme en ligne VIAN WebApp. Les utilisateurs peuvent en retour télécharger des projets depuis VIAN WebApp afin de l'adapter selon leurs propres intérêts. En outre, VIAN WebApp relie les projets aux galeries correspondantes disponibles sur la *Timeline of Historical Film Colors*. Enfin, l'application ColorMania est devenue une extension pour les visiteurs de l'exposition *Color Mania* au Fotomuseum Winterthur (figure 5).

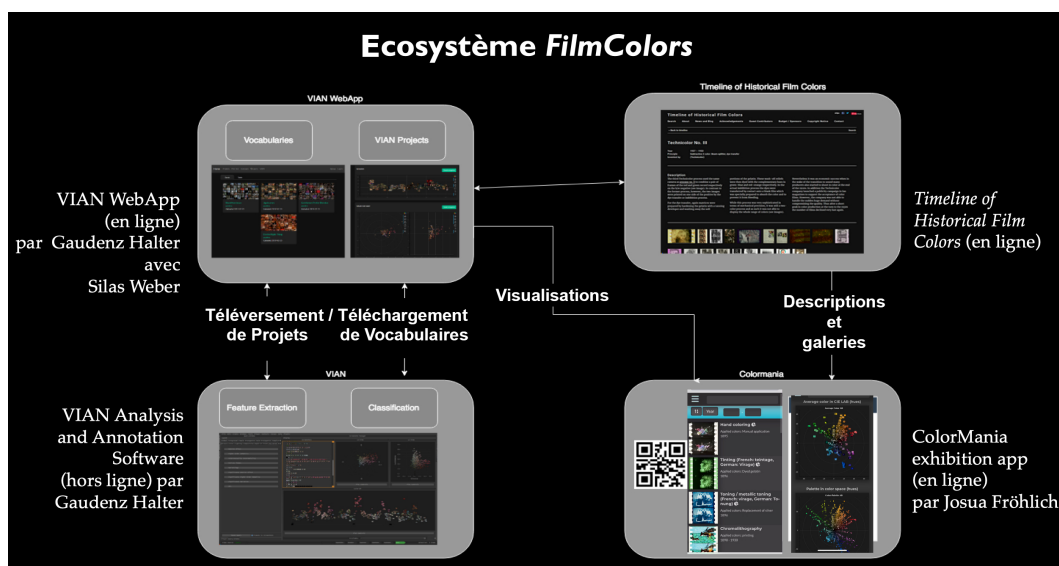


Figure 5. VIAN fait de plus en plus partie d'un écosystème composé de l'outil d'analyse et d'annotation VIAN et de la plateforme en ligne VIAN WebApp, connectés à la *Timeline of Historical Film Colors* et à l'application de l'exposition *ColorMania*.

Toutes les bases de données FileMaker ont été exportées vers l'architecture de bases de données VIAN WebApp. Chacune des bases de données individuelles est ainsi affichée sur l'application Web.

La **base de données de corpus** comprend les résultats répertoriés pour chaque film individuel, c'est-à-dire au niveau méso. Elle correspond à la page du projet sur VIAN WebApp. Pour chacun des champs d'Analysis DB, la base de données de corpus répertorie toutes les occurrences dans ce film, y compris une liste de tous les commentaires laissés dans les champs destinés aux « Remarques » avec l'ID du segment correspondant. Cette vue d'ensemble fournit une empreinte instantanée pour chaque film et constitue la base d'hypothèses donnant lieu à de nouvelles études.

Sous un angle différent, les résultats sont affichés dans la **base de données de glossaire** par concept individuel, avec de nombreux filtres personnalisés pour les périodes, l'attribution du corpus, les genres, le pays, etc. Suivant cette perspective, il est possible d'obtenir des informations instantanées sur la prédominance d'un trait narratif, d'un trait sémantique ou esthétique, d'un motif ou d'un lieu au niveau macro, triés par fréquence. La base de données du glossaire correspond à la page des concepts dans VIAN WebApp.

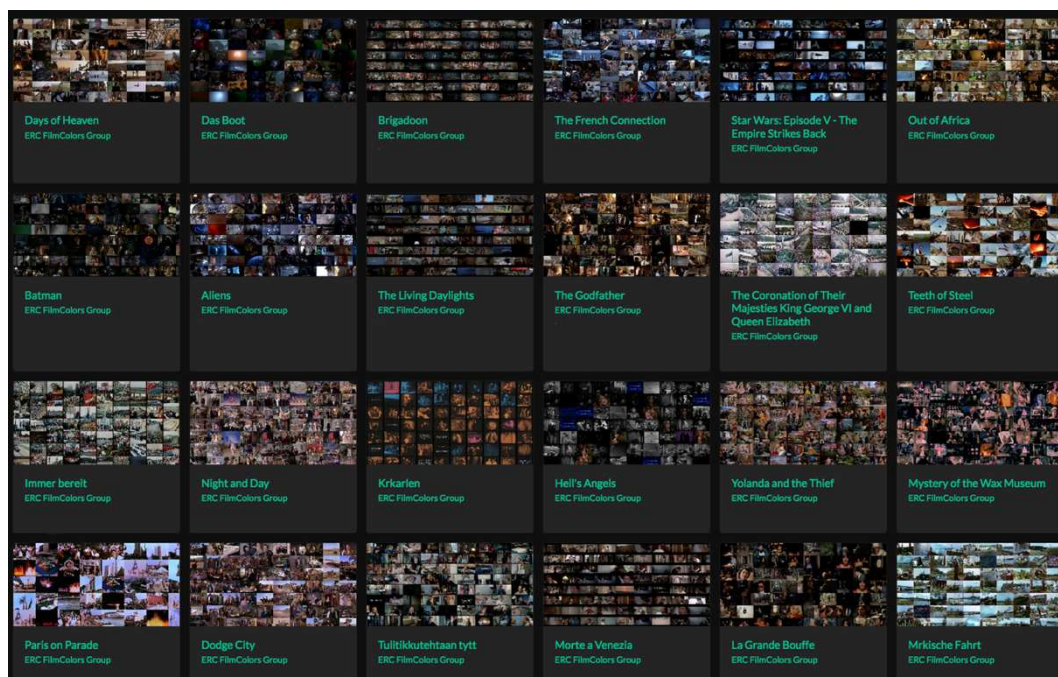


Figure 6. Vue d'ensemble du projet dans VIAN WebApp avec plus de 400 films analysés. Chaque projet comporte sa propre page avec des analyses et des visualisations détaillées – voir capture vidéo d'écran pour PARIS, TEXAS, <https://vimeo.com/396548709>

Enfin, la **base de données d'évaluation** elle-même, où tous ces résultats sont stockés, permet également l'analyse diachronique au niveau macro sur l'ensemble du corpus. Elle fournit des diagrammes d'évolution au fil du temps pour la période qui couvre les cent premières années de l'histoire du cinéma, de 1895 à 1995. Par exemple, l'évolution de certains styles d'éclairage (tels la lumière colorée, l'éclairage ambiant ou l'éclairage mixte), de types de profondeur de champ, des compositions complexes de l'image en plusieurs couches, est rendu visible en un coup d'œil. Ces tendances servent alors de fondement pour des hypothèses qui doivent être étudiées en détail dans la base de données d'analyse. Il est également essentiel de garder à l'esprit – comme nous le verrons plus en détail dans la section centrée sur les problèmes rencontrés (voir Section 3.1) – que ces résultats ne constituent pas nécessairement des faits concrets. Ils donnent un aperçu des tendances qui doivent ensuite être étudiées plus en détail et testées par le biais d'autres moyens d'évaluation. Or, les résultats dépassent de loin les approches traditionnelles et souvent anecdotiques qui amenaient beaucoup moins de preuves quant aux évolutions historiques.

Conformément à l’objectif du projet de recherche, à savoir l’identification des corrélations entre les innovations techniques et l’esthétique, l’identification des motifs diachroniques à l’aide des outils des humanités numériques constitue la base la plus importante pour de nouvelles perspectives allant au-delà des résultats précédents. En plus de montrer les liens importants entre la technologie et les formes stylistiques, elle révèle également que de tels liens causals ont souvent été surestimés dans le passé alors que les influences culturelles ou inter-médiatiques ont été largement négligées. Comme l’ont montré plusieurs études récentes, l’intégration minutieuse des visualisations colorimétriques dans la réflexion sur l’esthétique matérielle des films offre de nouvelles perspectives sur l’esthétique matérielle au cinéma – voir, par exemple, une étude récente des films expérimentaux en couleur réalisés par Len Lye dans les années 1930 (Flueckiger 2019) (figure 7).

Les captures d’écran elles-mêmes sont devenues une partie de plus en plus importante de l’étude. Initialement, seulement trois captures d’écran représentatives étaient intégrées dans la première version de la **base de données de glossaire (Glossary DB)** afin d’illustrer les concepts. Lorsqu’il s’est avéré que la base de données de glossaire était également un moyen parfait pour organiser les captures d’écran, son architecture a dû être entièrement modifiée pour intégrer les captures d’écran de manière dynamique via une deuxième base de données, la **base de données d’images du glossaire (Glossary Images DB)**, les deux bases de données étant connectées à la base de données de corpus. Avec cette architecture de bases de données, il est devenu possible d’écrire des scripts pour des portails et de trier les images en fonction des périodes, des attributions de corpus ou de la typologie. Pour VIAN WebApp, il est obligatoire de sélectionner un échantillon d’images les plus représentatives pour les utilisateurs externes, ce qui se fait en attribuant des priorités aux captures d’écran afin de les répertorier. L’architecture de bases de données FileMaker permet désormais aux captures d’écran d’être intégrées dans la base de données de corpus pour fournir une sélection des formes d’expression les plus importantes utilisant la couleur dans un film donné.

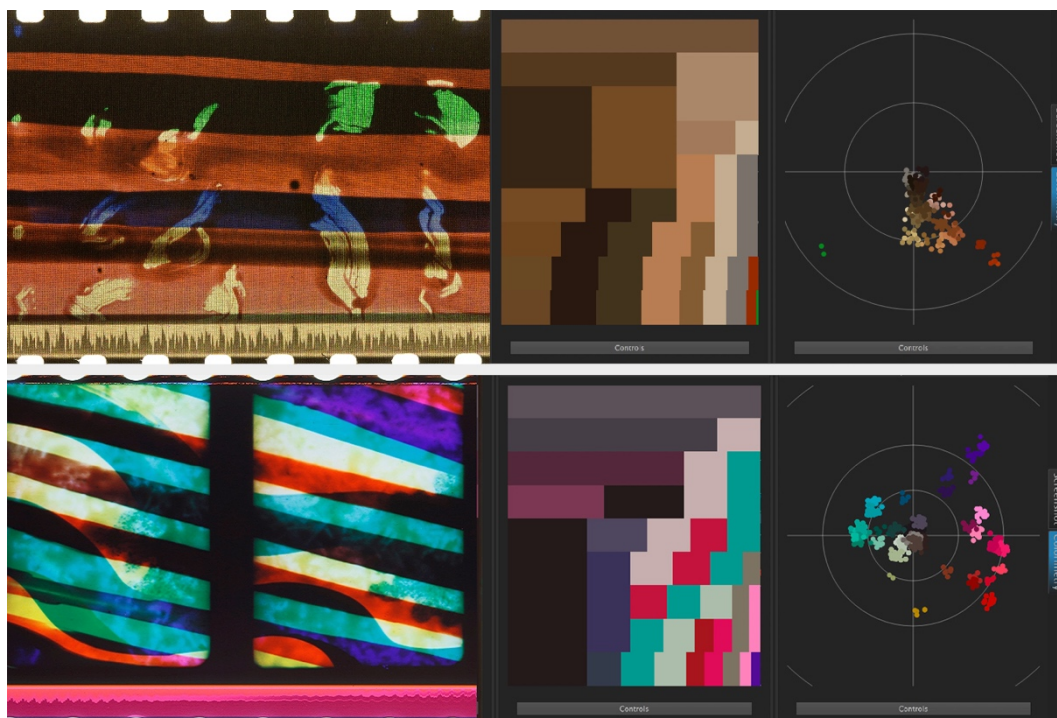


Figure 7. La comparaison de l'analyse colorimétrique réalisée sur VIAN offre des perspectives sur l'esthétique matérielle des films en couleurs historiques : les teintes brunâtres tamisées de Dufaycolor dans *A COLOUR BOX* (Grande-Bretagne 1935, Len Lye) vs la saturation et la séparation des couleurs élevées dans une copie Gasparcolor de *COLOUR FLIGHT* (Grande-Bretagne 1937, Len Lye)

Voir les galeries sur la *Timeline of Historical Film Colors*

https://filmcolors.org/filter/?_sft_ubercategory=lye&post_types=gallery

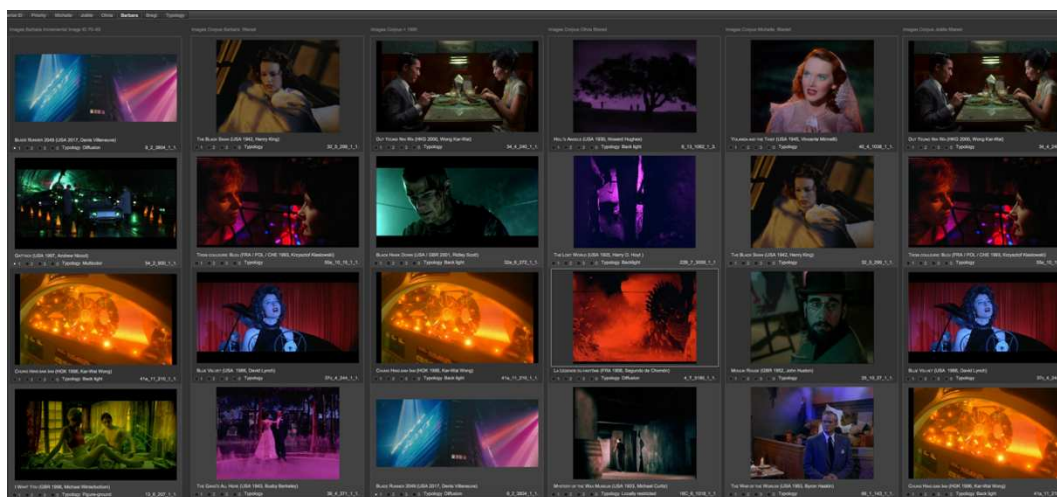


Figure 8. Base de données de glossaire : extrait d'une sélection représentative du concept de « lumière colorée », répertoriée selon les périodes, les sous-corpus ou les typologies.

[illegible]

Toutes les facettes de l'analyse ont été davantage intégrées dans l'écosystème global (voir la figure 5 et les Sections 4 à 9), ce qui a guidé le développement de la plateforme de *crowdsourcing* et la mise en œuvre des différents aspects de notre recherche.

Comme mentionné plus haut, une quantité considérable de données et de captures d'écran ont émergé pendant ce travail. Celles-ci ont été assemblées dans une base de données d'analyse principale qui aurait dû

être hébergée sur le serveur FileMaker de l'UZH. Or, il s'est avéré que ce serveur n'était pas configuré pour une tâche aussi exigeante, ce qui signifiait que toutes les captures d'écran devaient être exportées, réduites en taille et réimportées. Cela semblait être une tâche insurmontable pour FileMaker en raison de la nomenclature non standardisée pour les fichiers d'image. Par conséquent, nous avons finalement décidé d'exporter chaque capture d'écran dans un dossier individuel avec une nomenclature définie comprenant l'ID de l'élément du film, l'ID d'image et l'ID du plan – elles ont ensuite été traitées en externe par Gaudenz Halter et réimportées automatiquement avec un script dans FileMaker. Le traitement impliquait de redimensionner et de compresser les images ainsi que de trouver les horodatages des captures d'écran à l'intérieur du film. Puisque les captures d'écran exportées ne correspondaient pas exactement au contenu de leur cadre correspondant, en raison du redimensionnement et de la compression réalisés plus tôt dans le pipeline, le meilleur cadre correspondant a été déterminé en appliquant l'erreur quadratique moyenne. Ce mécanisme permet également d'importer des captures d'écran existantes dans un projet VIAN et de les attribuer automatiquement bon emplacement dans la vidéo.

Des problèmes similaires se sont produits lors de l'évaluation des données. Les scripts dans FileMaker utilisés pour assembler les données en résumés pour chaque film sont devenus trop complexes et le processus incroyablement lent et vulnérable. Même les spécialistes internes de FileMaker et les experts externes ne parvenaient pas à proposer de solutions. Par conséquent, nous nous sommes tournés vers un flux opérationnel similaire pour exporter toutes les données et les traiter en externe dans plusieurs scripts Python organisés dans un pipeline. Dans un premier temps, nous avons dû migrer l'ensemble des données exportées de FileMaker vers l'architecture de bases de données VIAN WebApp. Nous avons ensuite calculé les fréquences des mots-clés par film et associé des matrices de corrélation parmi ces mots-clés. Cette phase a été suivie d'un ensemble d'étapes successivement réalisées, conçues pour enrichir l'ensemble des données existantes avec des fonctionnalités de couleur numériques, y compris le calcul des histogrammes de couleurs, des palettes de couleurs et des valeurs de couleur moyennes pour chaque segment et

capture d'écran basée sur la séparation figure-fond. Enfin, les fréquences des mots-clés par film ont été réimportées dans FileMaker pour évaluation. Chaque membre de l'équipe ayant reçu sa propre couche d'analyse organisée en fonction de ses préférences et de ses intérêts, de nombreuses incohérences sont apparues, affectant des détails tels que l'orthographe, les concepts locaux et globaux, etc. La situation s'est encore compliquée par le fait que le glossaire a été élargi au fil du temps et, par conséquent, qu'il y avait des incohérences internes affectant le lien entre la logique tripartite de la taxonomie dans le glossaire, comprenant les classes, les champs et les concepts, et l'organisation des valeurs dans les champs de la base de données.

3.2 Leçons tirées

Pour collecter des données cohérentes et significatives, il est nécessaire de former les utilisateurs autant que possible et d'illustrer les concepts avec des captures d'écran exemplaires. La base de données de glossaire contient donc un champ prioritaire pour sélectionner les captures d'écran les plus informatives et les plus claires pour chaque concept – idéalement au moins six captures d'écran de périodes différentes pour chacun d'eux. Comme indiqué précédemment, ces captures d'écran sont immédiatement disponibles sur la page des concepts de VIAN WebApp (figure 10), de sorte que les utilisateurs aient une très bonne idée de ce quoi se réfère chacun des mots-clés. Ces catalogues de captures d'écran peuvent être enrichis avec des courts clips vidéo, tirés du corpus, dans les cas où le mouvement ou d'autres types de changement dans le temps sont jugés primordiaux pour le concept en question.

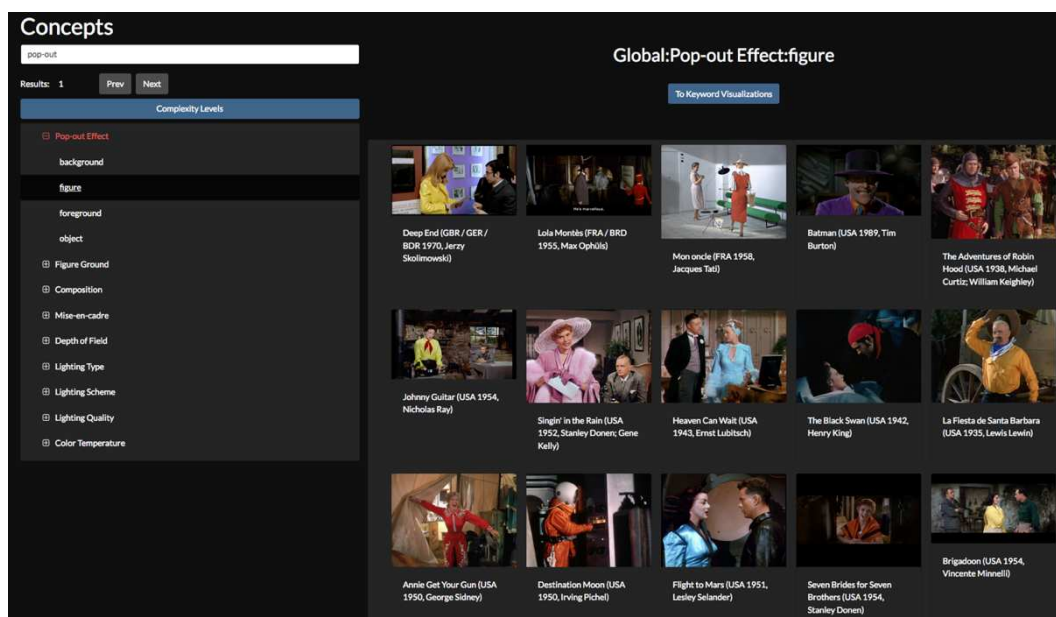


Figure 10. La page des concepts de VIAN WebApp : collection de captures d'écran représentatives pour le concept « effet pop-out » attribué à un sujet humain (par opposition aux objets).

4 Développement d'une plateforme d'annotation visuelle, d'analyse et de visualisation

Les annotations manuelles réalisées en 2016 et 2017 avec la combinaison d'ELAN et les bases de données FM, ont permis de développer un ensemble d'outils pour des analyses de couleurs semi-automatiques et automatiques ainsi que pour la visualisation des résultats. Ces outils emploient la vision d'ordinateur et l'apprentissage profond afin de produire des résultats significatifs (Flueckiger et al. 2017 ; Flueckiger 2017 ; Flueckiger/Halter 2018 ; Halter et al. 2019).

Les outils d'annotation vidéo ont été parmi les premières approches utilisées afin d'appliquer des méthodes numériques pour segmenter et annoter des films – voir, par exemple, les travaux de Gruber, Wurm et Kropf (2009), les études dans Giunti (2010, 2014) et une évaluation détaillée par Melgar et al. (2017).

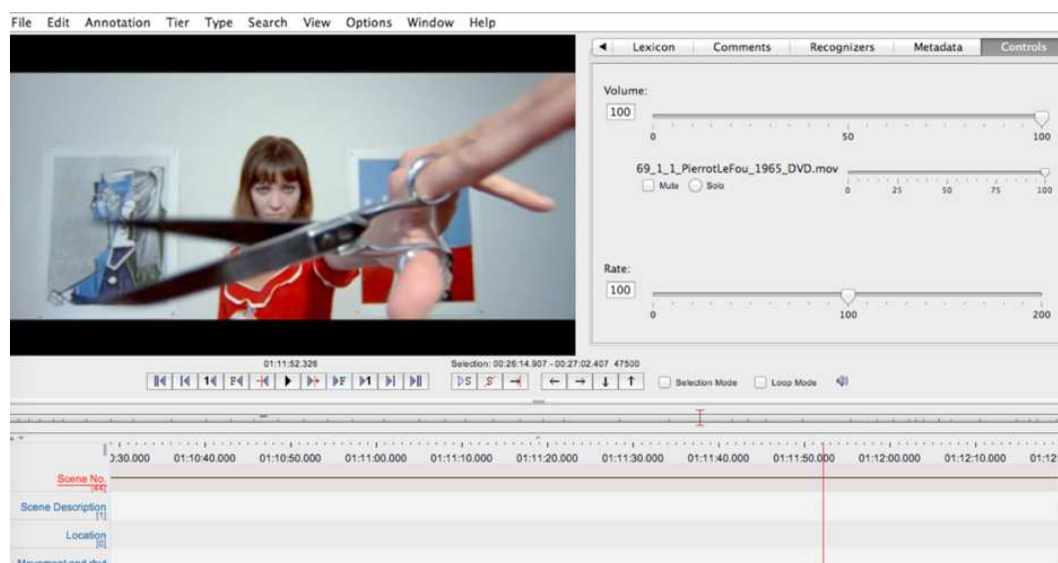


Figure 11. L'interface du système d'annotation vidéo ELAN et un modèle de segmentation et d'annotation, PIERROT LE FOU (France 1965, Jean-Luc Godard)

En 2016, nous avons mené des recherches approfondies sur tous les outils disponibles, dont beaucoup ne fonctionnaient plus sur les nouveaux systèmes d'exploitation, en raison d'une interruption de financement (voir Flueckiger 2017). Nous avons finalement décidé d'utiliser ELAN (figure 11), un outil d'annotation vidéo qui offre une grande variété d'options et qui est très sophistiqué ; or cet outil a été développé par l'Institut Max Planck de psycholinguistique à Nijmegen en mettant l'accent sur l'analyse du langage.

Pour surmonter les limites de cette approche et pour se concentrer davantage sur la perspective des formes d'expression visuelles, le nouveau système VIAN d'annotation visuelle des vidéos a été développé par Gaudenz Halter depuis 2017. Il intègre plusieurs couches de segmentation et d'annotation vidéo, ainsi que des méthodes avancées d'analyse et de visualisation des couleurs des films tout en étant bien adapté pour la classification à grande échelle du contenu filmique.



Figure 12. Couche de segmentation VIAN avec gestionnaire de captures d'écran, SEDMIKRÁSKY [LES PETITES MARGUERITES] (Tchécoslovaquie 1966, Vera Chytilová)

VIAN est un logiciel d'annotation de films basé sur plusieurs niveaux qui met l'accent sur les aspects visuels du style des films et leur esthétique en termes de couleurs, permettant à l'utilisateur d'effectuer des tâches générales d'annotation ainsi que l'analyse numérique du matériau filmique. VIAN a été développé non seulement pour fournir des données à l'outil de *crowdsourcing* VIAN WebApp, qui rassemble toute notre analyse en un seul logiciel, mais aussi pour une utilisation flexible dans d'autres projets de recherche liés à l'analyse de film sur des sujets connexes. Il se compose essentiellement de plusieurs ingrédients cruciaux : gestion des captures d'écran, classification par nomenclatures, un ensemble d'outils pour l'analyse des couleurs et des visualisations ; pour un aperçu élémentaire, voir les divers tutoriels pour l'outil d'annotation VIAN : <https://vimeo.com/user/70756694/folder/1220854>

À notre connaissance, les outils d'annotation précédents n'implémentent pas de *système de gestion des captures d'écran*, de sorte que les captures d'écran doivent généralement être exportées et gérées par l'utilisateur dans le système de fichiers, une tâche difficile qui est naturellement sujette à erreur avec un nombre important de captures d'écran. Néanmoins, comme indiqué précédemment, les captures d'écran jouent un rôle clé dans l'évaluation visuelle des films. Elles sont ainsi devenues une forme d'annotation cruciale qui peut être créée dans VIAN. Outre les captures d'écran, VIAN fournit également des segments temporels et des annotations graphiques vectorielles. Ces dernières

désignent des annotations qui peuvent être dessinées directement à l'écran – elles incluent actuellement des ellipses, des rectangles, des images, du texte et des dessins à main levée.

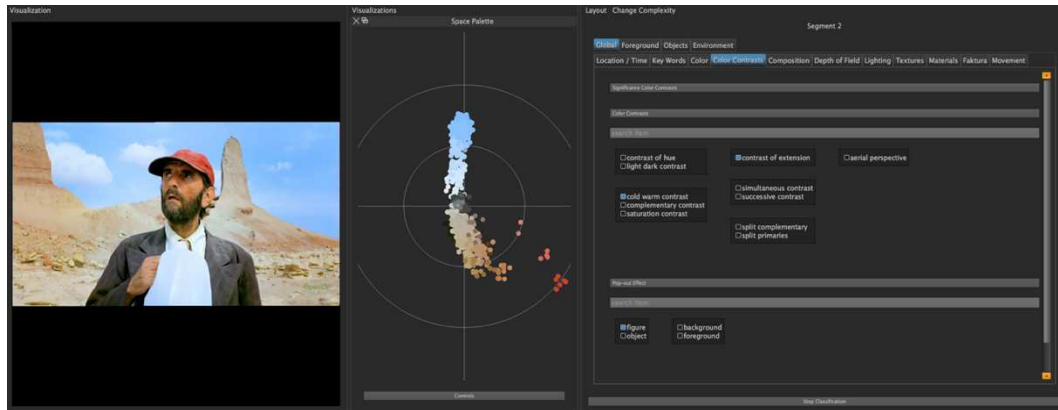


Figure 13. L'interface graphique d'analyse VIAN contient le vocabulaire contrôlé, développé lors des analyses manuelles et défini dans la base de données de glossaire ainsi que sur la page des concepts de VIAN WebApp : PARIS, TEXAS (Allemagne/France 1984, Wim Wenders).

Comme mentionné précédemment (Section 2), notre projet inclut la classification d'un grand nombre de segments à l'aide de plus de 1200 concepts analytiques et théoriques via FileMaker. En ce qui concerne l'outil de *crowdsourcing* VIAN WebApp, cette fonctionnalité a été implémentée ainsi. Contrairement à de nombreux autres logiciels d'annotation de film, VIAN fait une distinction claire entre l'annotation basée sur le langage naturel et la *classification* basée sur une nomenclature établie et testée dans l'analyse manuelle du corpus. Les descriptions sont effectuées en tapant simplement l'annotation respective dans le segment temporel ou en tant qu'annotation graphique vectorielle sur l'écran. Cependant, cette dernière est réalisée au sein du système de classification de VIAN. Une fois que l'utilisateur a créé une ou plusieurs annotations composées de captures d'écran, de segments temporels ou d'annotations graphiques vectorielles, elles peuvent être classées selon la nomenclature définie dans le glossaire. VIAN permet également à l'utilisateur de définir l'entité conceptuelle, lesdits *objets de classification* qui sont indexés explicitement sous un ou plusieurs termes. Par exemple, le concept « saturé » pourrait cibler l'objet de classification « protagoniste masculin » et « protagoniste féminin ». Les fonctions de couleur peuvent être extraites pour une annotation afin de créer des visualisations qui offrent un aperçu du contexte colorimétrique des captures d'écran, des segments temporels et

des régions dans le cadre. En outre, VIAN calcule automatiquement plusieurs mesures colorimétriques de manière équitablement espacée pour que le film complet affiche directement les caractéristiques des couleurs les plus importantes lorsqu'on parcourt ou visionne une vidéo.

Nous avons opté pour le langage Python et avons mis l'accent sur l'extensibilité, afin que les chercheurs puissent facilement développer les fonctionnalités de VIAN pour répondre aux besoins de leurs propres questions de recherche ou l'utiliser comme une API Python.

4.1 Problèmes

Le développement de VIAN a été un processus itératif continu avec l'équipe de recherche. De toute évidence, un grand nombre de questions de design émergent au cours d'un tel processus, en particulier lorsque le nombre d'exigences et de tâches est aussi important que dans le cas de notre recherche sur les couleurs des films. Bien entendu, de nombreuses questions sont liées à l'architecture du logiciel et à l'implémentation des outils, mais nous avons également constaté que le développement d'un logiciel facile à utiliser peut en réalité s'avérer être compliqué. L'une des principales difficultés liées à l'architecture de VIAN a été le développement d'un logiciel qui réponde aux besoins très spécifiques de notre projet de recherche tout en restant suffisamment générique pour être utilisé dans le cadre d'autres projets et sujets de recherche.

Une autre difficulté était liée au stockage efficace des données. La plupart des outils d'annotation utilisent un format de fichier lisible par l'utilisateur tel que XML ou JSON afin de stocker en permanence les données générées. L'avantage de tels formats est que les données peuvent être facilement lues, même sans l'outil source à portée de main, améliorant ainsi l'interopérabilité. Néanmoins, les données numériques générées et exploitées par VIAN ont dû être stockées dans un format de fichier auquel on pouvait accéder plus rapidement. Au cours du développement, nous avons essayé plusieurs approches. Nous avons commencé avec un simple fichier JSON. Une fois que les données numériques sont devenues trop volumineuses, nous avons migré vers une base de données SQLite. Or, cette approche n'était pas assez évolutive ; ainsi, nous avons finalement mis en place un système hybride utilisant un fichier JSON lisible par

l'utilisateur pour les annotations et la structure du projet, et un fichier HDF5 optimisé pour les données numériques.

4.2 Leçons tirées

Nous avons constaté que l'aspect le plus important du développement est une boucle de rétroaction courte entre le développeur et les utilisateurs, les spécialistes en cinéma, les étudiants et d'autres chercheurs. Puisqu'il existe une vaste palette de méthodes statistiques et analytiques qui pourraient être implémentées dans VIAN, il s'est avéré qu'un développement centré sur l'utilisateur est plus favorable que l'implémentation d'un large éventail de fonctionnalités possibles. Ainsi, nous avons essayé de créer une base architecturale solide et de rester générique autant que possible sans introduire trop de complexité.

5 Segmentation temporelle, extraction et organisation des captures d'écran

Les approches de segmentation temporelle (*parsing*) des films varient considérablement en fonction des intérêts du chercheur :

Ils peuvent être analysés de manière significative en une hiérarchie qui propose des unités dans les unités dans les unités [...] Au sein de cette hiérarchie, certaines unités possèdent la stature psychologique des événements. Autrement dit, les spectateurs les considèrent comme ayant un début, un milieu et une fin, délimitées par des frontières qui sont souvent marquées par les changements de temps et de lieu, et qui forment des segments séparables à l'intérieur du flux audiovisuel en cours.

(Cutting, Brunick, Candan 2012 : 1)

Ce qui pourrait sembler relativement simple s'avère plutôt complexe, surtout quand on considère les outils de segmentation (semi-)automatique (Hahn 2009). Les ambiguïtés augmentent lorsque nous définissons des unités temporelles en fonction de la cohérence de leurs schémas de couleurs, à savoir l'objectif de l'étude des couleurs des films. Même si l'angle de prise de vue varie ou si la caméra se déplace via des travellings ou des mouvements de grue, les couleurs qui dominent dans une séquence peuvent varier de manière significative quoique continue. Par conséquent, il s'avère difficile d'identifier les segments temporels de manière cohérente.

Certains modèles de montage – comme le montage parallèle – nécessitent des sous-segmentations prenant en compte à la fois les frontières des événements et les conflits entre les unités temporelles. Alors que les films muets avec leurs intertitres et/ou les segments uniformément teintés signalent souvent leur organisation structurale d'une manière assez claire, les films les plus récents comprennent des séquences qui se chevauchent de manière plus fluide. D'autre part, le système de continuité du cinéma classique hollywoodien a établi un certain nombre de marques d'énonciations qui communiquent les changements de séquence ou les ellipses temporelles, comme les fondus enchaînés, les fondus au noir ou les volets.

La segmentation temporelle d'un film par des observateurs humains – et en particulier par ceux qui ont été formés en tant que spécialistes du cinéma ou les étudiants avancés en études cinématographiques – prend en considération tous ces indices différents et historiquement établis, même si la tâche est principalement liée à la dimension de la couleur du film. Les membres de l'équipe ont identifié en moyenne entre cinquante et soixante-dix unités par film avec des schémas de couleurs suffisamment cohérents.

Afin d'accélérer ce processus fastidieux, VIAN propose une fonctionnalité d'auto-segmentation qui calcule une segmentation temporelle au moyen d'un regroupement aggloméré d'histogrammes de couleurs régulièrement espacés. Le résultat peut ensuite être affiné par l'utilisateur à l'aide de l'outil de fusion et de découpe sur le ligne de temps de VIAN.



Figure 14. Comparaison de la segmentation manuelle (en haut) par rapport à quatre types de segmentation temporelle automatique avec 30 à 60 segments dans *UNE FEMME EST UNE FEMME* (France 1961, Jean-Luc Godard)

Comme expliqué plus haut (voir Section 4), une sélection raisonnée de captures d'écran est un outil heuristique essentiel pour réduire la complexité du flux temporel vidéo en sélectionnant les moments les plus pertinents. Ainsi, il est devenu impératif d'établir un processus rapide et

flexible non seulement pour extraire les captures d'écran avec une commande simple mais aussi pour les organiser instantanément en fonction de la segmentation temporelle du film dans des bacs individuels et avec une nomenclature cohérente. Dans ELAN, chaque extraction de capture d'écran nécessitait plusieurs étapes impliquant cinq à douze commandes, tandis que les fichiers d'image devaient également être nommés manuellement et le format d'image défini. En revanche, VIAN traite les captures d'écran comme un type d'annotation intégral, et leur création et leur gestion constituent donc des fonctionnalités clés. Les captures d'écran sont créées avec un raccourci clavier simple, affichées de plusieurs manières dans VIAN, y compris l'alignement temporel dans la ligne de temps, et regroupées par segmentation dans le gestionnaire de captures d'écran.

5.1 Problèmes

De nombreuses solutions logicielles pour l'annotation des vidéos ont été établies dans le passé pour répondre aux besoins élémentaires. Pourtant de réels avancées sont encore nécessaires afin de développer des outils plus sophistiqués qui répondent à des exigences plus complexes. Autrement dit, le diable se cache dans les détails. Un cadre très fin qui intègre plusieurs types de lecteurs pour différentes fonctions de zoom arrière est un moyen puissant pour segmenter temporellement les films, annoter verbalement les unités qui en résultent et extraire des captures d'écran. Dans quelle mesure le lecteur intégré traite-t-il divers codecs et formats d'image ? Quelles options offre-t-il pour ajuster les segmentations et ajouter des sous-segmentations pour des entités discontinues tel le montage parallèle ? Comment empêche-t-il le chevauchement des segments ou, inversement, comment les permet-il ? Quelles sont les options pour corriger les segmentations existantes ?

La segmentation temporelle automatique fournit des résultats étonnamment bons qui, dans certains cas, rivalisent avec les approches humaines pour subdiviser le flux vidéo en morceaux cohérents. Néanmoins, l'auto-segmentation semblait être beaucoup plus finement détaillée s'agissant des séquences sombres, et certains segments étaient trop longs, notamment par rapport à la longueur moyenne des segments.

5.2 Leçons tirées

Afin de développer un système solide d'annotation vidéo, un retour constant de la part d'utilisateurs expérimentés est nécessaire. Pour la prochaine étape d'auto-segmentation, nous envisageons de prendre en compte les indices musicaux et le traitement du son (pour une approche très originale de combinaison image-dialogue dans l'analyse des films avec des outils numériques, voir, par exemple, Burghardt/Kao/Wolff (2016)). Très souvent, le lancement ou l'interruption d'une musique diégétique indique un changement de lieu ou de temps. Le traitement du son peut également être associé à certains lieux ou indices temporels.

En outre, les soi-disant marques d'énonciation telles le fondu au noir (ou au blanc), les fondus enchaînés ou les intertitres devraient être incorporées dans le système de règles pour l'analyse des unités. Les écarts importants quant à la longueur des segments obtenue par rapport à la moyenne devraient contraindre le système à traiter de nouveau ces morceaux extrêmement longs ou courts.

6 Séparation figure-fond

Très tôt dans le projet, un outil de séparation figure-fond a été mis en place (Flueckiger et al. 2017) afin d'extraire les personnages de l'arrière-plan à l'aide d'une technique de segmentation sémantique d'apprentissage profond (Long, Shelhamer/Darrell 2015, Zhao et al. 2016). Cette approche découle de l'idée d'attribuer une étiquette à chaque pixel d'image et indique l'objet représenté le plus probable. Elle vise à étudier l'esthétique de l'attribution des couleurs à travers les costumes et la scénographie en conjonction avec d'autres paramètres de la mise en scène. Dans le projet, la méthode a été constamment améliorée en termes de vitesse et de performance, fournissant la base pour tous les autres outils d'analyse des couleurs – *LAB plots* (voir les figures 18 et 21, etc.), les tracés *Color_dT* (voir les figures 20 et 22) – qui considèrent les personnages indépendamment de l'arrière-plan.

L'esthétique de la séparation figure-fond a beaucoup varié au cours de l'histoire du cinéma, en fonction de nombreux facteurs tels les procédés technologiques des couleurs, la photographie et la mise en scène – y

compris l'éclairage, les emplacements dans le champ et la matérialité des costumes, des objets et des environnements – ainsi que des notions de goût et de normes professionnelles. Par exemple, une forte séparation figure-fond était un moyen stylistique typique pour améliorer la lisibilité immédiate dans le contexte dudit *système de continuité*, établi dans le cinéma classique hollywoodien à partir du milieu des années 1920.

Dans ce contexte de production, il existait souvent une hiérarchie selon laquelle les couleurs les plus attrayantes visuellement étaient attribuées à la star féminine et les différences de couleur entre les personnages et les arrière-plans étaient minimisées s'agissant des personnages secondaires. La saturation est principalement attribuée aux personnages féminins, tandis que les personnages masculins ne portaient des costumes colorés que quand ils incarnaient des rôles déterminés à l'intérieur des normes culturelles, par la distance historique – par exemple, dans le cas des rois ou des uniformes –, par certains milieux comme celui de l'industrie du divertissement, par l'altérité culturelle comme l'exotisme ou par des traits de personnalité individuels comme la *queerness* ou des attitudes non-conformistes (voir Bohn 2000, Vänskä 2017), voire par les conventions de genre. La tendance pour une forte séparation figure-fond s'observe à nouveau dans les contextes émergents des premiers films d'auteur en couleur dans des productions européennes ou, par exemple, japonaises, au style sobre et moderniste.

Afin d'étudier ces changements stylistiques et culturellement justifiés de manière claire, nous avons établi une typologie qui a pris en compte les dimensions suivantes : fort contre faible, silhouettes, inversion figure-fond, et séparation par teinte, par saturation ou par clarté. Par inversion figure-fond, nous entendons des relations dans lesquelles l'arrière-plan est soit plus saturé, soit plus lumineux que l'avant-plan. Ces distinctions sont ensuite devenues les concepts sous-jacents des visualisations issues du pipeline de séparation figure-fond.

En ce qui concerne le système d'annotation et de classification expliqué précédemment, VIAN permet à l'utilisateur de définir les *objets de classification* pour exprimer une entité conceptuelle d'intérêt, dans ce cas « figure » et « fond ». VIAN utilise une segmentation sémantique

basée sur l'apprentissage profond afin d'interpréter le contenu d'un cadre (figure 15). Le produit d'une telle segmentation est une image en niveaux de gris, dans laquelle chaque pixel de l'image-source est assigné à une valeur de gris, une sorte d'*étiquette*, qui correspond à des objets définis sur lesquels le modèle a été formé. VIAN permet désormais d'attribuer un ensemble d'étiquettes à chaque objet de classification, créant un lien sémantique entre le contenu du cadre et la classification effectuée par le chercheur. Les études de la culture visuelle par Arnold et Tilton ont également appliqué la reconnaissance d'image avec des outils d'apprentissage profond (Arnold et Tilton 2020a et b).

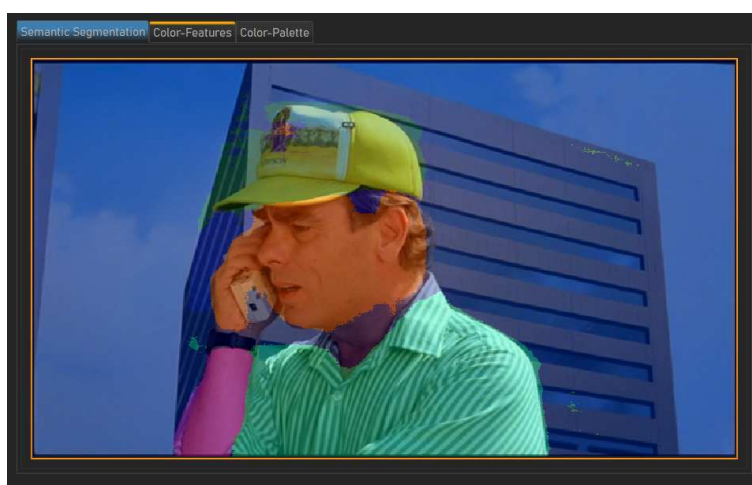


Figure 15. Segmentation sémantique réalisée dans VIAN

Les résultats du procédé figure-fond sont très significatifs, en particulier lorsqu'ils sont combinés avec la visualisation *Color_dt* (voir figure 20). Une empreinte digitale instantanée de l'évolution esthétique au sein d'un film émerge lorsqu'on compare les relations variables entre l'attribution des couleurs aux personnages et à l'environnement au cours du déroulement du récit. Comme cela sera développé dans la Section 7, *Analyses et visualisations colorimétriques*, les types de visualisation qui en résultent diffèrent profondément des visualisations établies. La cartographie des résultats soulève encore quelques questions en termes de mise à l'échelle. Par exemple, nous avons constaté que les êtres humains perçoivent les niveaux de saturation attribués aux personnages comme étant plus élevés lorsque le reste de l'image est moins saturé. Cette différence qui ne peut pas encore être rendue avec précision par le biais de nos méthodes actuelles de visualisation et de colorimétrie.



Figure 16. Séparation figure–fond dans VIAN, JIGOKUMON (Japon 1953, Teinosuke Kinugasa)

6.1 Problèmes

Nous nous attendions à ce que ce travail soit très exigeant, mais il s’est avéré que – puisqu’il s’agit de l’une des tâches les plus importantes de la conduite automatique – les méthodes d’apprentissage profond sont actuellement dans un état très dynamique, en particulier lorsqu’il s’agit d’extraire les personnages de l’arrière-plan. YOLO (Redmon et al. 2015) a été le premier logiciel adapté à la reconnaissance d’objets. Il a fourni des résultats très fiables pour l’identification des êtres humains, contrairement à celle d’autres objets qui ont souvent été mal interprétés, en particulier lorsqu’ils étaient partiellement cachés ou coupés sur les bords du cadre.

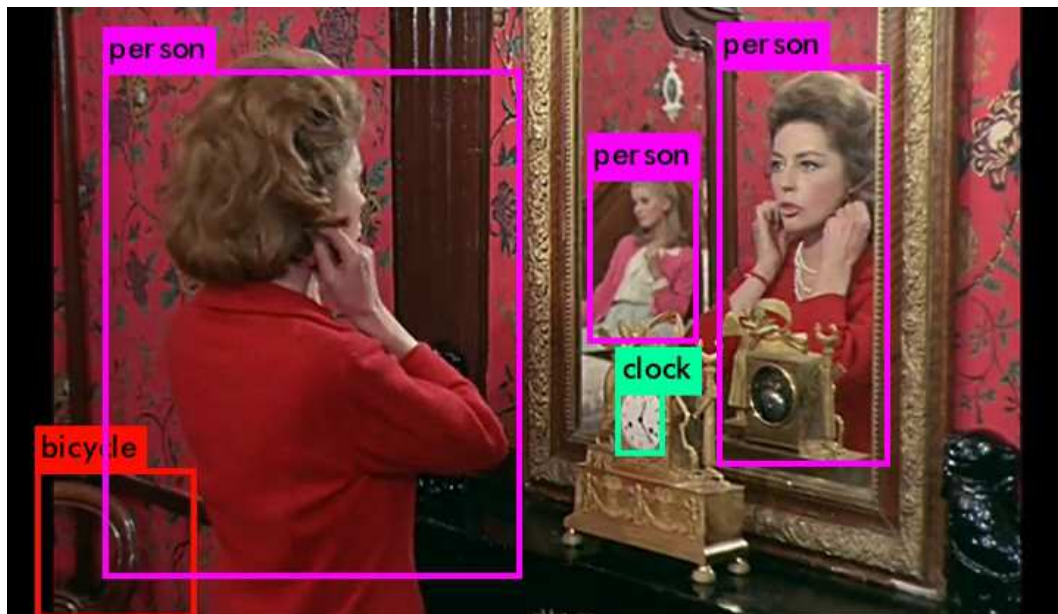


Figure 17. Identification des objets dans YOLO : alors que les personnes sont identifiées de manière systématique, des erreurs se produisent lorsque des objets sont partiellement cachés ou fragmentés. LES PARAPLUIES DE CHERBOURG (France 1964, Jacques Demy)

YOLO a été combiné avec GrabCut (Rother, Kolmogorov et Blake 2004). GrabCut fonctionne de la manière suivante : l'utilisateur dessine initialement un rectangle autour de l'objet assigné au premier plan, et GrabCut s'efforcera ensuite de segmenter directement le cadre et de retourner le résultat. L'utilisateur peut alors optimiser le résultat en marquant des régions qui n'ont pas été correctement identifiées à l'aide de traits. L'exécution manuelle de ce processus pour chaque image n'aurait pas été possible en raison des contraintes de temps ; ainsi, nous avons utilisé YOLO, un réseau neuronal de reconnaissance d'objets, pour dessiner le cadre de sélection initial et les traits. Or, ce procédé n'était pas suffisamment évolué pour nos besoins et nécessitait une quantité importante de ressources et de temps. Ainsi, nous avons décidé d'utiliser un réseau convolutif d'apprentissage profond afin de réaliser la segmentation des pixels directement avec la segmentation sémantique (Long, Shelhamer et Darrell 2015) plutôt qu'avec une approche basée sur YOLO et GrabCut.

6.2 Leçons tirées

Une approche collaborative et interdisciplinaire reliant des niveaux élevés d'expertise dans les domaines de l'analyse esthétique et de l'informatique s'est avérée nécessaire pour l'élaboration d'un procédé d'analyse et de visualisation qui respecte les deux champs et les relie de manière convaincante. Bien qu'une telle déclaration puisse sembler banale, l'échange réel entre les différentes disciplines s'est révélé être plus exigeant que prévu, nécessitant des ajustements constants des deux côtés impliqués. Les spécialistes du domaine des sciences humaines doivent être en mesure de comprendre les exigences de l'informatique et de décrire la tâche de manière très formalisée. Inversement, les scientifiques doivent être capables d'intégrer la subtilité des concepts esthétiques afin de comprendre pourquoi des détails mineurs peuvent, de manière tout à fait inattendue, avoir un impact significatif sur les résultats. Le procédé qui en émerge devrait produire des visualisations qui respectent les exigences rigoureuses de la science informatique tout en considérant l'accessibilité instantanée pour les observateurs humains couplée à une connaissance des distinctions esthétiques.

7 Analyses colorimétriques et visualisations

Les approches antérieures pour une visualisation des schémas de couleurs avaient une portée étonnamment étroite et n'étaient pas assez sophistiquées pour rendre justice aux subtilités esthétiques du traitement de la couleur au cinéma.

Les outils qui sont actuellement disponibles pour générer des schémas de couleurs appliquent souvent des k-moyennes (Brodbeck 2011, Rosebrock 2014) et sont ainsi bornés à représenter seulement un ensemble fixe de teintes. En revanche, les schémas de couleurs dans VIAN sont extraits de sorte qu'ils correspondent à la distribution spatiale et peuvent être traités en fonction du type d'analyse des couleurs qui est nécessaire pour un film donné. Certains films appliquent des teintes très distinctes à leurs schémas de couleurs, tandis que d'autres ont recours à des décalages infimes afin de montrer les développements dans les relations entre les personnages. Les schémas de couleurs peuvent exprimer l'état intérieur d'un personnage, l'étape de son développement personnel ou sa relation à d'autres personnages ou à un environnement donné ; ils peuvent également être liés à des normes de goût, des milieux sociaux ou des conventions culturelles. Les repères sociopolitiques indiquent l'attachement d'un personnage à une certaine classe ou à une fonction sociale de manière prédéfinie en termes sociaux ou culturels, comme on peut le voir par exemple dans un vêtement comme les uniformes.

Ainsi, nous avons envisagé d'entrée de jeu une voie différente qui permettrait une approche flexible pour affiner les schémas de couleurs de sorte qu'ils correspondent au style spécifique d'un film donné. Une deuxième exigence fondamentale a été la représentation de la distribution spatiale des couleurs, afin de pouvoir afficher de manière instantanée l'attribution quantitative des couleurs dans une image ou un segment temporel. Troisièmement, nous avons cherché des moyens de visualiser les schémas de couleurs afin d'indiquer leur développement au fil de la durée d'un film en fonction de la segmentation temporelle réalisée dans le pipeline.

Les représentations temporelles typiques telles que les codes-barres ou les mosaïques donnent des aperçus plénoptiques des films (voir Heftberger

2016 ; Stutz 2016 ; Olesen 2017 ; Flueckiger 2017), mais ne représentent pas les fluctuations et les relations fines qui sont essentielles pour une étude esthétique approfondie. Frederick Brodbeck a organisé les schémas de couleurs en cercles pour donner un aperçu de ce qu'il a nommé une empreinte digitale du schéma des couleurs d'un film (Brodbeck 2011). Les Z-projections constituent un aspect principal des visualisations de Kevin L. Ferguson (Ferguson 2013, 2016) ; il a également proposé une approche volumétrique comme un moyen de visualiser la couleur d'un film dans son ensemble sur l'axe temporel en 3D (Ferguson 2015). James E. Cutting et son équipe à l'Université Cornell ont mis au point de nombreuses méthodes pour visualiser des films, parmi lesquelles un code-barres qui implémente des schémas de couleurs allant des couleurs chaudes aux couleurs froides (Cutting 2016). À partir de 2013, Lev Manovich (2013a et b) et son laboratoire Software Studies ont appliqué une gamme de visualisations aux films de Dziga Vertov pour le projet de recherche d'Adelheid Heftberger (Heftberger 2015, 2016, 2018), dont certaines sont basées sur ImagePlot et ImageJ, précédemment utilisées pour la visualisation des œuvres d'art (Manovich 2012 ; Reyes-García 2014, 2017). Initialement lancé pour la recherche biomédicale (Ross 2007), ImageJ a depuis été employé par plusieurs chercheurs pour l'analyse des films (Olesen 2016 ; Heftberger 2016 ; voir également plusieurs chapitres dans Acland/Hoyt 2016). Casey et Williams (2014) ont comparé des segments temporels dans des films sur la base d'histogrammes visualisés dans une matrice de similarité.

Comme expliqué dans Halter et al. (2019 : 126), l'équipe a déterminé un ensemble d'exigences pour les visualisations :

- 1) représenter des impressions visuelles fidèles à la perception humaine
- 2) représenter des nuances esthétiques subtiles distinctes pour la figure et le fond, et
- 3) visualiser les films aux niveaux micro (captures d'écran, segment temporel), méso (un film à part entière) et macro (un corpus des films),

tout en restant interactives et flexibles pour qu'elles puissent être adaptées en fonction des intérêts individuels du chercheur.

Par conséquent, comme nous avons expliqué dans des articles précédents (Flueckiger 2011, 2017), les relations des couleurs doivent être cartographiées en un espace de couleurs perceptuellement uniforme afin de fournir des visualisations qui correspondent à la vision humaine. Dans VIAN, les captures d'écran et les schémas de couleurs sont ainsi transformés en l'espace colorimétrique CIE L^*a^*b (appelé LAB dans le présent article) nécessaire pour une représentation signifiante de la distribution des couleurs dans un film donné. Contrairement à la plupart des représentations visuelles établies telles que les tracés d'image, Z-projections, les palettes de couleurs ou les codes-barres, une représentation visuelle en un espace colorimétrique perceptuellement uniforme prend en considération la nature relationnelle des couleurs vis-à-vis du système visuel. Les tracés de chromaticité et de clarté donnent un aperçu de la distribution des couleurs au sein d'un film (voir la figure 18).

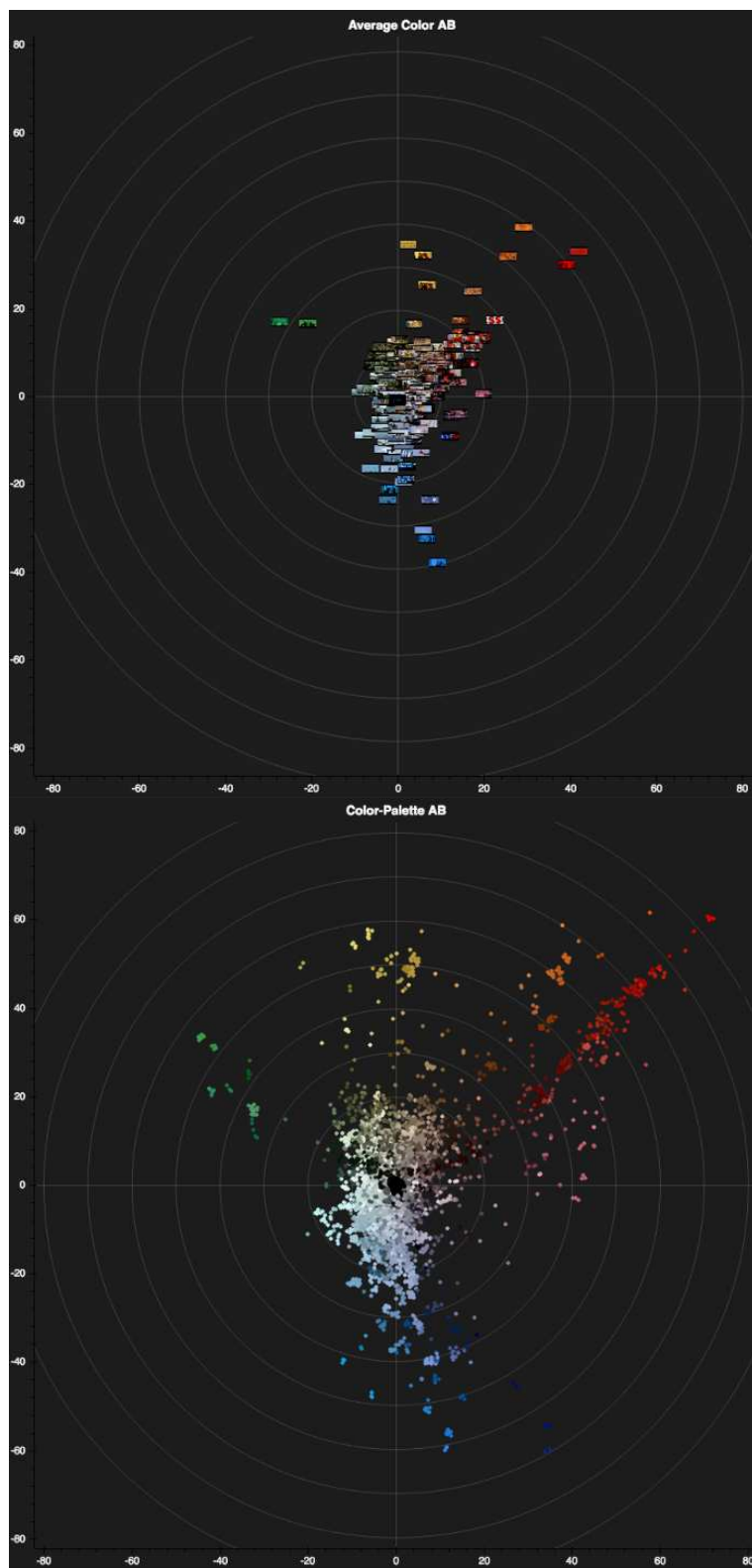


Figure 18. Tracés de chromaticité dans CIE L*a*b (LAB) pour PIERROT LE FOU (France 1965, Jean-Luc Godard), graphique d'images (en haut) vs graphique de points de palette (en bas). La comparaison montre que la chrominance s'étend beaucoup plus loin dans les tracés de points de palette sans les effets de pondération produits par la représentation d'images entières.

Bien que l'utilisation des schémas de couleurs de visualisation dans un groupe de couleurs triés par fréquence soit généralement bien établie et offre une vue d'ensemble satisfaisante, ces schémas sont souvent difficiles à comparer et masquent la manière dont la palette a été assemblée pendant le processus de regroupement. Afin de comparer la distribution des couleurs par rapport à la perception humaine, le schéma de couleurs est affiché dans l'espace colorimétrique LAB sous forme de tracés de points de palette. Grâce à cette méthode, les changements subtils et les contrastes de couleur au sein de la chrominance ou la teinte entre les différents groupes de couleur deviennent directement visibles (voir figure 19, à droite). La palette arborescente (figure 19, en haut et au milieu) devrait aider l'utilisateur à comprendre dans quel groupe final les couleurs de l'image-source ont été fusionnées. À cette fin, les palettes sont empilées les unes sur les autres dans des étapes séparées de fusion, correspondant aux niveaux croissants de granularité, et les groupes de couleurs sont triés à l'intérieur de la palette selon l'ordre résultant du regroupement. De cette manière, toutes les couleurs fusionnées dans un groupe sont visualisées juste en dessous de l'image.

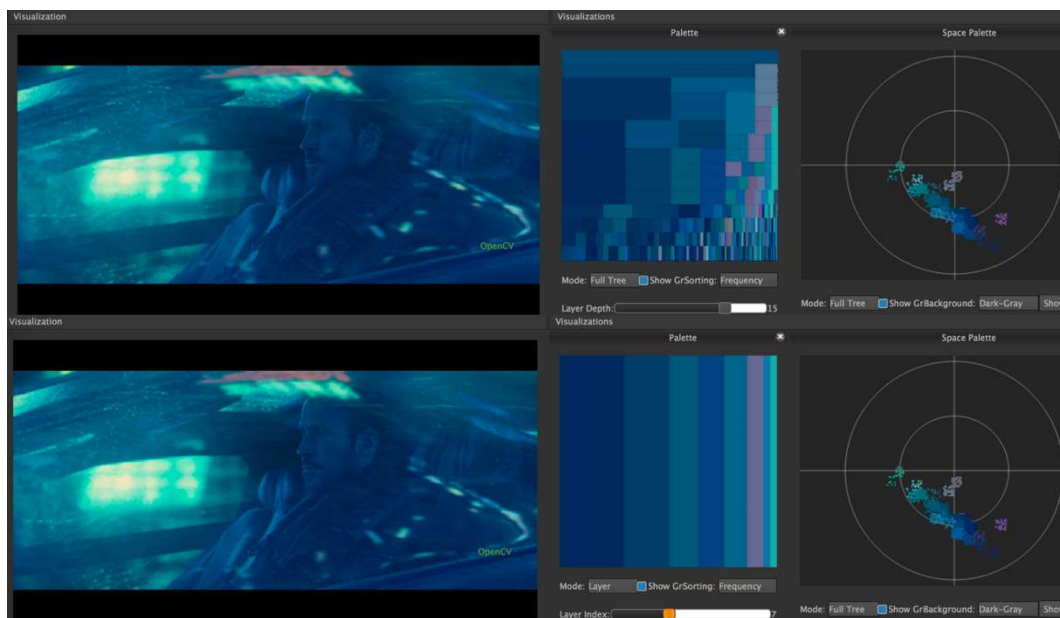


Figure 19. Palettes de couleur pour des plans séparés : diagramme arborescent aux niveaux de détail croissants (au-dessus et au milieu) et LAB (à droite) ; une sélection de sept teintes pour la palette de calques triées par fréquence (en dessous, au milieu). BLADE RUNNER 2049 (États-Unis 2017, Denis Villeneuve) – voir les méthodes de visualisation interactive dans la capture vidéo d'écran <https://vimeo.com/299804415>

L'application des méthodes computationnelles dans l'analyse des couleurs, telles que le regroupement ou les statistiques sur les images brutes d'un film, n'est souvent pas possible en raison de la quantité énorme des données. Il est ainsi souvent nécessaire d'extraire des vecteurs des traits qui représentent de manière adéquate le contenu au moyen d'un histogramme de couleurs régulièrement utilisé dans VIAN. Néanmoins, les espaces colorimétriques sont tridimensionnels, tout comme leurs histogrammes de couleurs, ce qui rend la visualisation de ces derniers difficile. Une approche naïve consisterait à visualiser les histogrammes sous forme de nuages de points dans un espace tridimensionnel ; or, cette méthode ne donne pas une bonne comparabilité. Par conséquent, nous avons développé une représentation d'un histogramme de couleurs qui ressemble à un diagramme à barres en triant les couleurs de l'histogramme tridimensionnel dans une liste unidimensionnelle à l'aide d'une courbe de remplissage d'espace, à savoir la courbe de Hilbert. Intuitivement, cette courbe de remplissage d'espace décrit un chemin par lequel tout point d'un espace donné est visité – dans notre cas les bacs de l'histogramme de couleurs tridimensionnel. En démêlant cette courbe, nous pouvons aligner les bacs de l'histogramme de couleurs tridimensionnel en une rangée unidimensionnelle. Nous utilisons des courbes de Hilbert car il a été démontré que ce type de courbe de remplissage d'espace préserve bien la localité spécifique – dans notre cas, cela signifie que les bacs de couleur qui sont proches les uns des autres dans l'histogramme tridimensionnel seront également proches dans le diagramme à barres unidimensionnel démêlé.

En revanche, *Color_dT* est une méthode avancée de visualisation du développement des couleurs au fil de la durée d'un film au niveau méso, lié à son déroulement temporel – en se concentrant de nouveau de manière indépendante sur la figure, le fond et les captures d'écran dans leur entièreté. Il est actuellement mis en œuvre pour la saturation, la teinte et la chrominance ou les contrastes clair-obscur, mais pourrait également inclure un contraste froid-chaud. Les variations dans les relations figure-fond deviennent instantanément évidentes, comme d'ailleurs les développements globaux concernant les événements narratifs au cours d'un film.

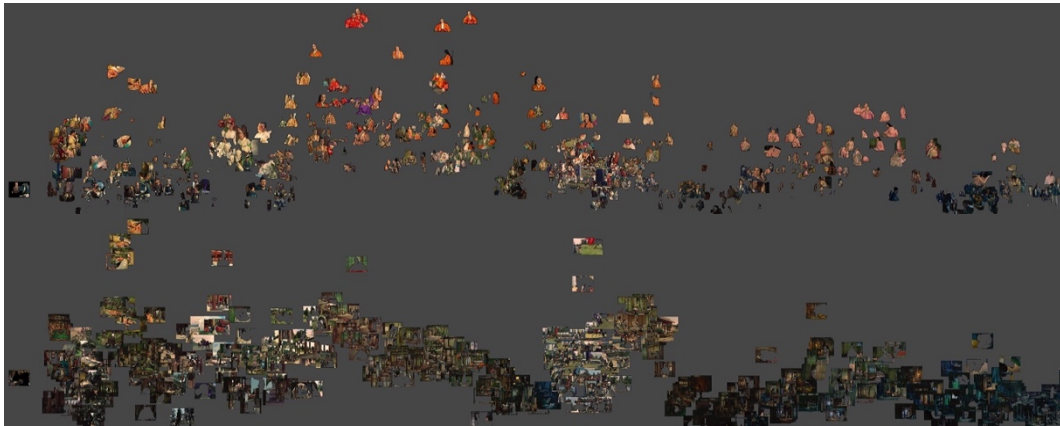


Figure 20. Tracés *Color_dT* de l'évolution de la saturation (axe y) en fonction du temps (axe x) pour la figure (en haut) et le fond (en bas) dans JIGOKUMON (Japon 1953, Teinosuke Kinugasa) ; visualisation par Noyan Evirgen pour ERC Advanced Grant *FilmColors*

JIGOKUMON, l'un des premiers films en couleurs japonais à être filmés avec le procédé chromogène Eastmancolor, fournit un exemple significatif. Dans la première moitié du film, on note une séparation prononcée figure-fond avec les personnages, en particulier la femme qui devient l'objet d'un intérêt romantique, se démarquant dans des kimonos saturés aux motifs colorés devant des arrière-plans tamisés. Au milieu du film, une péripétie se produit lors d'une course de chevaux dans laquelle les deux adversaires masculins s'affrontent. Cette séquence se déroule en plein jour avec des couleurs qui s'opposent entre l'arrière-plan et le premier plan. Après ce tournant, la tragédie s'installe avec un traitement de la couleur nettement différent et une mise en scène caractérisée par des scènes sombres, éclairées en *low-key*, à savoir un type d'éclairage qui par sa nature même réduit la séparation figure-fond.

Avec les premières couleurs appliquées telles que la teinture et le virage, les tracés de chromaticité LAB semblent résolument différents en raison des schémas de couleur principalement monochromes. Comme le montre la comparaison entre L'INHUMAINE (France 1923, Marcel L'Herbier) et DAS CABINET DES DR. CALIGARI [LE CABINET DU DR. CALIGARI] (Allemagne 1919, Robert Wiene), la numérisation de L'INHUMAINE est sensiblement différente quant à la distribution détachée de la chrominance – pas de continuité entre le centre et les niveaux supérieurs –, ce qui pourrait indiquer l'existence de problèmes dans la gestion numérique des couleurs.

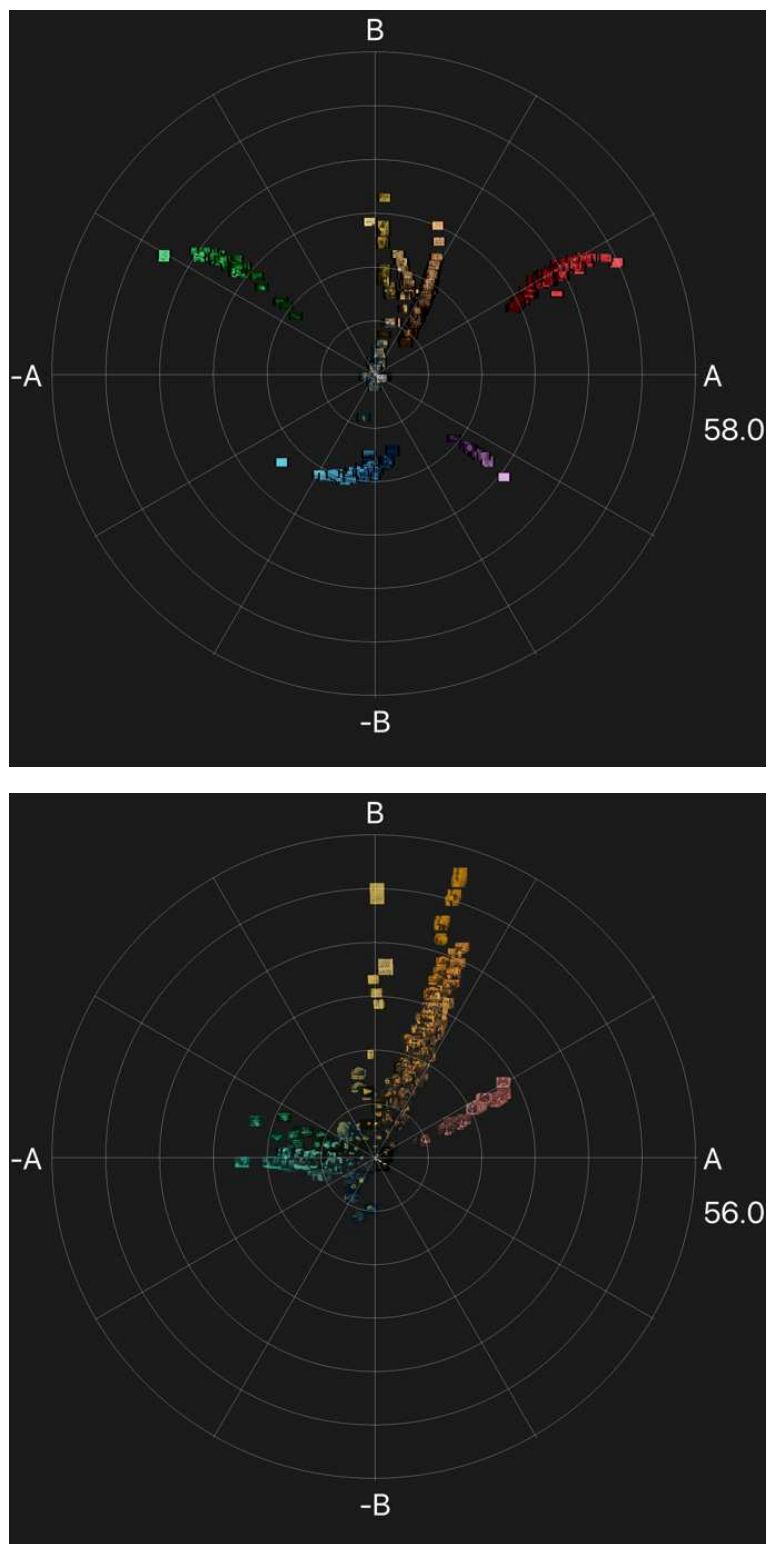


Figure 21. Tracés de chromaticité dans l'espace CIE L*a*b pour les films teintés L'INHUMAINE (France 1923, Marcel L'Herbier) (en haut) et DAS CABINET DES DR. CALIGARI (Allemagne 1919, Robert Wiene) (en bas)



Figure 22. Tracés *Color_dT* de l'évolution de la saturation (axe y) en fonction du temps (axe x) pour le film teinté *L'INHUMAINE* (France 1923, Marcel L'Herbier).

Les corrélations entre les concepts sont affichées de deux manières distinctes. L'outil de fonctionnalités permet aux utilisateurs de sélectionner les concepts dans le menu. Puis, l'occurrence de ces concepts est affichée en fonction du temps, par rapport aux segments où ils apparaissent. Lorsqu'il est présenté en combinaison avec les captures d'écran sélectionnées comme représentatives, ce type de visualisation crée instantanément une base pour établir et tester des hypothèses.

La corrélation entre différents mots-clés au sein d'un projet ou à l'échelle d'un corpus peut être étudiée à l'aide des tracés matriciels de cooccurrence indiquant la fréquence à laquelle chacune des combinaisons de mots-clés se produit à l'intérieur du champ de référence.

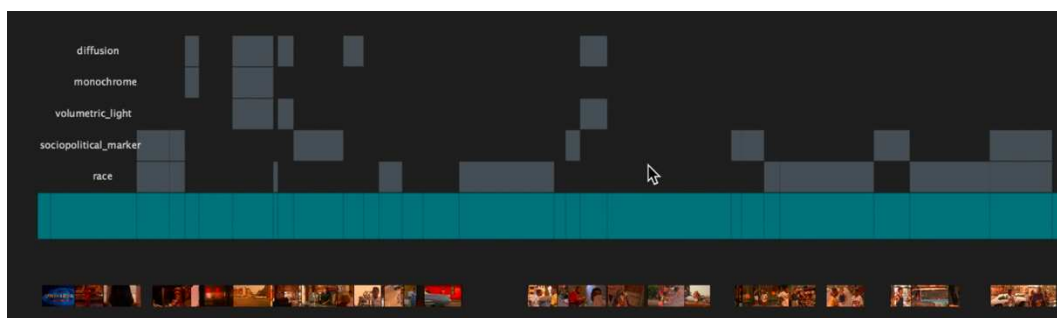


Figure 23. L'outil de fonctionnalités VIAN montre les cooccurrences des concepts organisées sur l'axe temporel par rapport à la segmentation temporelle dans *DO THE RIGHT THING* (États-Unis 1989, Spike Lee) – voir la capture vidéo d'écran <https://vimeo.com/292861139>

Dans la capture vidéo d'écran (voir le lien ci-dessus), l'outil de fonctionnalités est testé avec le film de Spike Lee, *DO THE RIGHT THING*. Plusieurs traits typiques du film ont été sélectionnés dans cette visualisation. Outre le leitmotiv qui établit le spectre rouge dominant du film et l'associe à la thématique de la chaleur (*heat*) au double sens du mot – renvoyant à la fois à la température élevée et à la montée des tensions raciales –, l'esthétique du film est imprégnée par une séparation entre la sphère privée et l'espace public. La première est souvent représentée, dans des espaces intérieurs, baignée d'une lumière diffuse ambiante, de nouveau associée à l'humidité brumeuse provoquée par la

chaleur, en tons rouges monochromes chauds avec des puits de lumière qui remplissent la pièce, le tout lié à une tradition romantique datant du début du XIX^e siècle. En revanche, l'esthétique du film dans sa représentation de l'espace public suit un style beaucoup plus sobre, lié aux traditions du réalisme social, avec une profondeur de champ étendue dans des plans filmés avec des objectifs à courte focale, qui montrent les personnages en relation les uns avec les autres et avec leur environnement. En termes de traitement de la couleur, le film emploie ce que nous appelons des repères sociopolitiques, des conventions culturellement établies qui dénotent certaines couches sociales ou des fonctions officielles. Par exemple, le protagoniste joué par Spike Lee lui-même est un livreur de pizzas et porte des vêtements aux couleurs du drapeau tricolore italien (blanc, vert et rouge). Ses rencontres sont également déterminées par la thématique centrale de la race, liée aux différents groupes ethniques – portoricains, jamaïcains, coréens, italiens et afro-américains –, chacun d'eux associé à un ensemble différent des teintes via des repères sociopolitiques. Un tel modèle d'esthétique et de signification des couleurs peut facilement être confirmé ou développé plus en profondeur via l'outil de fonctionnalités (voir figure 23).

7.1 Problèmes

Les tracés d'images sont des outils remarquables pour la visualisation lorsqu'un chercheur souhaite conserver la trace de la matière d'entrée. En zoomant sur les tracés, les utilisateurs peuvent regarder les captures d'écran et s'apercevoir où et pourquoi une certaine capture d'écran est présente dans le tracé et comment elle est liée au film. Cependant, comme les valeurs colorimétriques des tracés d'images sont calculées sur la base de la moyenne de la distribution des couleurs dans la capture d'écran, les schémas de couleurs monochromatiques déforment la visualisation en étant trop dominants. Les captures d'écran avec plus d'une teinte ou une multitude de teintes s'agrègent au centre de la visualisation LAB ou au bas d'un tracé *Color_dT*. Ainsi, ce type de visualisation est mieux adapté aux premières couleurs appliquées, comme la teinture et le virage avec leur distribution de couleur monochrome, ou pour les films caractérisés par un traitement de la couleur vif avec, la plupart du temps, une seule teinte

dominante par segment, à l'instar de SUSPIRIA ou le travail de Sławomir Idziak dont le style est reconnaissable par l'emploi fréquent d'un éclairage coloré et par des filtres d'objectif monochromes ou gradués.

7.2 Leçons tirées

Pour de nombreux films qui ne sont pas bien rendus en tracés d'images, nous avons conçu une solution alternative en nous passant du rendu de l'image complète et en séparant les valeurs de couleur individuelles (pour une comparaison, voir la figure 18). Autrement dit, au lieu d'utiliser les valeurs de couleur moyennes, nous avons calculé la palette de couleurs pour une capture d'écran donnée et l'avons visualisée dans le plan AB de l'espace colorimétrique LAB. Un effet de gigue est appliqué pour ajouter du bruit, ce qui rend la quantité d'une teinte spécifique visible au sein de l'espace colorimétrique. Ces visualisations des tracés de points de palette montrent maintenant les schémas de couleurs représentés par des points pour chacune des couleurs présentes dans une capture d'écran. Nous avons dû concevoir une méthode pour inclure le pourcentage spatial dans les tracés de points. Les tracés de points sont également devenus un moyen d'afficher les schémas de couleurs d'une manière différente par rapport à celle qui est possible avec les barres de couleurs classiques (voir figure 19). Des méthodes différentes pour la mise à l'échelle et la distribution des couleurs dans les visualisations sont proposées, telles que les fonctions zoom ou les ajustements de gamme. La rotation est cruciale pour les visualisations liées à l'axe L dans l'espace LAB afin de montrer la distribution des teintes d'une manière signifiante. Bien que les tracés de points de palette affichent la distribution des couleurs au sein d'un film de manière intuitive, ils ne prennent pas en compte l'incidence relative. Pour cette raison, un autre type de visualisation a été introduit : des cartes thermiques qui illustrent la distribution des couleurs en attribuant des niveaux de transparence qui correspondent à l'incidence (voir figure 24).

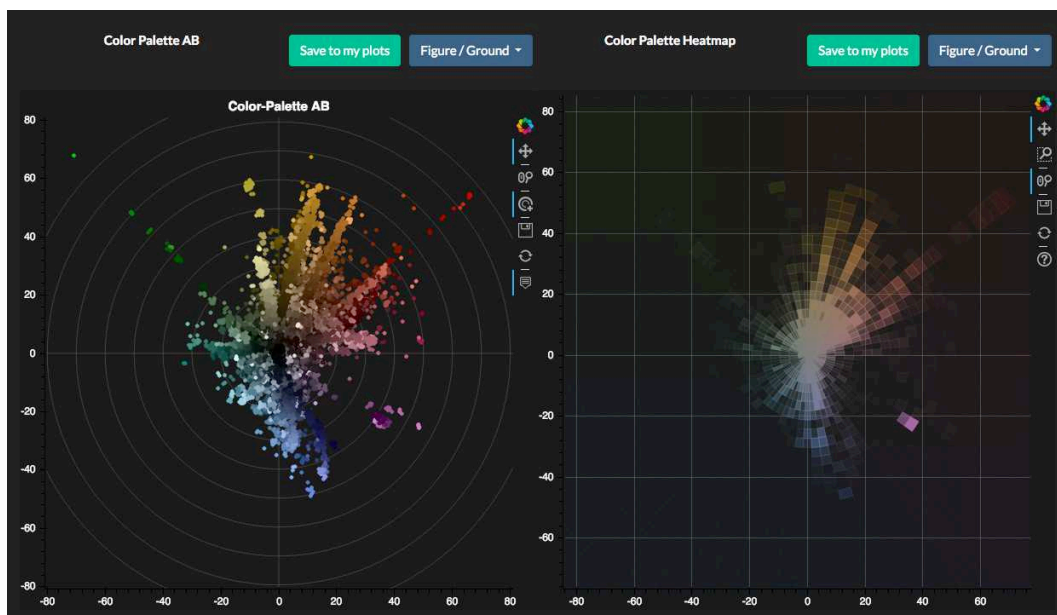


Figure 24. Tracés de points de palette (à gauche) vs cartes thermiques (à droite) : alors que les tracés de points de palette montrent l'occurrence d'une certaine couleur, les cartes thermiques indiquent l'incidence en observant les différents niveaux de transparence.

Requête VIAN WebApp : « dégoût »

8 Visualisations, concepts et corrélations au niveau macro

L'un des plus grands apports de notre étude est le vaste ensemble de données créées par l'équipe d'analyse. Comme indiqué plus haut (voir Section 3), cet ensemble comprend plus de 17000 segments avec plus de 170000 captures d'écran pour plus de 400 films, dont chacun est lié à l'analyse manuelle minutieuse et à l'annotation présentées en détail dans les sections 2 à 5 de cet article.

Une manière d'afficher le nombre de liens entre les différents mots-clés dans un ensemble de données aussi vaste et complexe consiste à suivre une approche de visualisation des réseaux. Chaque mot-clé est représenté comme un nœud et ses liens à d'autres mots-clés comme des arêtes. Plus ces mots-clés apparaissent dans le même segment, plus ils sont rapprochés dans le réseau à l'aide de l'algorithme de dessin graphique de force dirigée de Fruchterman-Reingold, fourni par la bibliothèque NetworkX Python (Hagberg/Schult/Swart 2008).

L'intégration de cet ensemble de données dans VIAN WebApp ouvre un large éventail d'opportunités pour des requêtes au niveau du segment, du film et du corpus, afin de combiner l'annotation manuelle avec toutes

les méthodes d'analyse colorimétrique et de visualisation élaborées dans les Sections 6 et 7. En adoptant une telle approche englobante, nous permettons aux utilisateurs de combiner les trois niveaux différents, depuis les niveaux micro (lecture attentive – par exemple, des captures d'écran ou des segments individuels) et méso pour des films individuels jusqu'aux niveaux macro (lecture à distance) de l'ensemble du corpus ou des sous-corpus sélectionnés. Des requêtes peuvent se réaliser sur ces corpus sélectionnés en utilisant n'importe quel concept lié aux aspects narratifs, aux états émotionnels ou aux motivations des personnages, aux thématiques, ainsi qu'à toutes les dimensions esthétiques et stylistiques évoquées dans la Section 2.

Deux requêtes de concepts sont montrées ici : la recherche de séquences de rêve pour les périodes 1895-1930 et 1930-1955, et la recherche de séquences nocturnes pour la première période. Lorsqu'on compare la visualisation des séquences de rêve des premières décennies de l'histoire du cinéma avec celles de la période qui va de 1930 à 1955, deux constats s'imposent : les séquences de rêve sont souvent marquées par des schémas de couleurs monochromes et, le plus souvent, la couleur dominante est le rouge. Dans la deuxième période (1930-1955), l'occurrence relativement importante du vert est liée de manière presque exclusive au WIZARD OF OZ, où le concept de rêve s'applique au récit principal du film (voir figure 25, en bas).

Les couleurs appliquées dans les films produits au cours des trois premières décennies de l'histoire du cinéma – comme la teinture et le virage avec leurs schémas de couleurs monochromes – ont suivi un ensemble de conventions, qui doivent ensuite être testées sur des films individuels ou sur une certaine période. Au vu des nombreuses ambiguïtés, les schémas de couleurs de chaque film et l'attribution des teintes à différents lieux, époques, courants narratifs, genres ou conventions de genre (*gender*) doivent être attentivement étudiés par les spécialistes en cinéma et les restaurateurs, afin de comprendre les principes directeurs pour chaque copie historique de film (Ledig/Ullmann 1988 ; Mazzanti 1998, 2009). Par exemple, LE CABINET DU DR. CALIGARI existe en cinq versions toniques et teintées différemment – voir la galerie sur la *Timeline of Historical Film Colors*, <https://filmcolors.org/galleries/das-cabinet-des->

dr-caligari-1919/ – mais la copie de référence de la version montrée à la première originale du film en Allemagne n'a pas encore été retrouvée (Wilkening 2014) (voir figure 9 pour la comparaison entre un DVD et une copie historique).

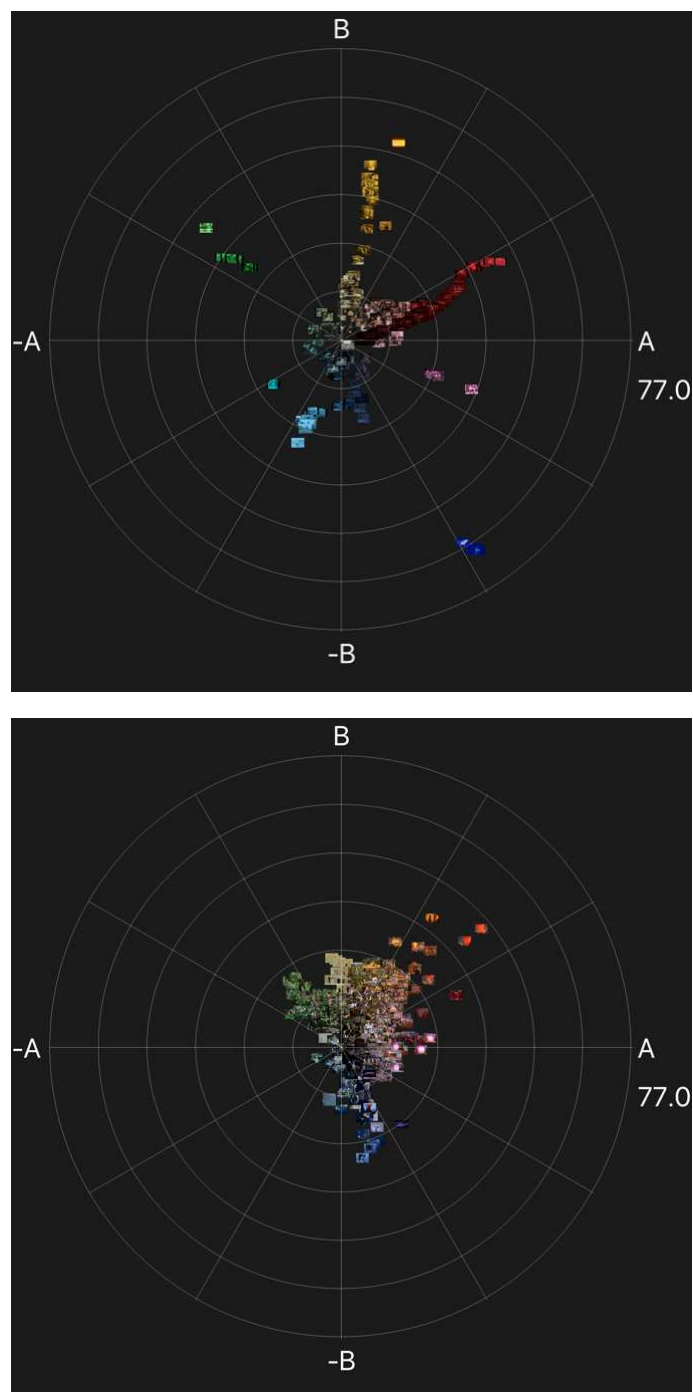


Figure 25. Des visualisations au niveau du corpus pour le concept narratif « rêve » : 1895-1930 (en haut) et 1930-1955 (en bas), tracés d'images AB.

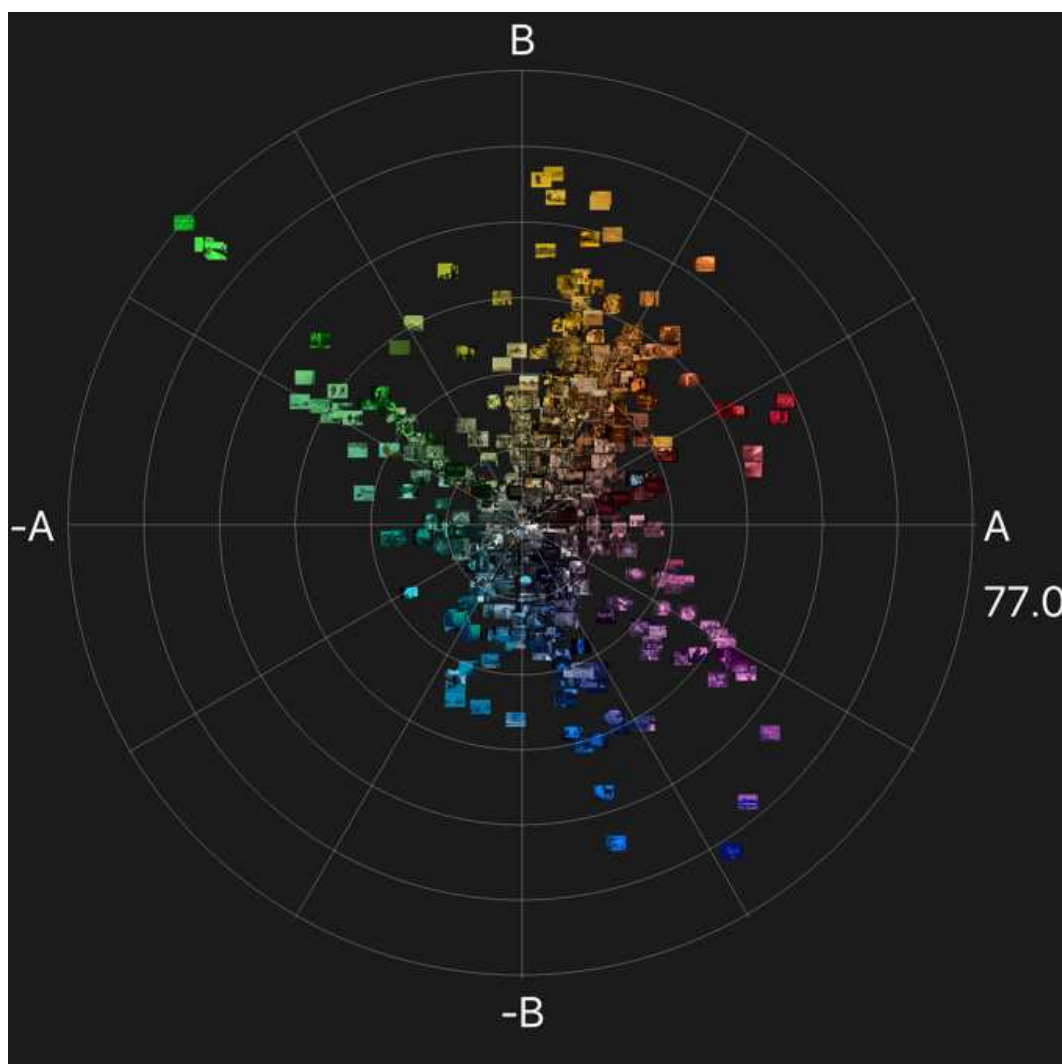


Figure 26. Visualisation du corpus de séquences nocturnes en extérieur dans des films datant de 1895 à 1930, à l'exclusion des scènes d'incendie.

L'une des associations les plus constantes entre une teinte spécifique et une certaine dimension narrative est la teinte bleue pour les séquences nocturnes en extérieur, étant donné que la vitesse limitée des premiers types de pellicule ne permettait pas aux séquences nocturnes d'être réellement tournées de nuit. Par conséquent, ces séquences devaient être marquées par des teintes typiques. Les visualisations montrent que le bleu est effectivement l'une des teintes dominantes, le vert étant presque aussi répandu que le bleu. L'ambre et le rouge sont la troisième gamme la plus dominante. L'ambre est souvent associé à la lumière au tungstène ou à la lumière des bougies dans des scènes en intérieur, de sorte que les segments qui contiennent des liens entre des scènes tournées en intérieur et en extérieur influencent en quelque sorte le résultat. En revanche, les

séquences éclairées par un feu sont généralement teintées en rouge et ont ainsi été éliminées de la requête (voir figure 26).

De manière générale, les tracés d'images LAB et les tracés de points de palette ont une valeur informative réduite au niveau du corpus, bien qu'ils soient utiles au niveau du film, en particulier lorsqu'ils sont affichés en version imprimée. Ils n'indiquent que des tendances qui doivent ensuite être confirmées en observant de manière plus approfondie les films et les segments où ils se produisent. À cette fin, toutes les visualisations de la page de requête de VIAN WebApp sont hautement interactives. Lorsqu'il survole les tracés, le chercheur peut voir les segments correspondants du film, y compris les captures d'écran et une description de la séquence. En outre, tous les segments et les films sont affichés avec les palettes de couleurs correspondantes sous forme de codes-barres bruts (voir la capture vidéo de la page de requête : <https://vimeo.com/402360042>).

Afin d'étudier le développement diachronique, une méthode supplémentaire pour la visualisation des corpus, nommée *Color_dY* a été implémentée ; elle prend en considération la distribution temporelle, au lieu de tracer une période sélectionnée en un aperçu dans LAB qui obscurcit les schémas de couleurs des films individuels.

Le film d'animation FANTASIA (États-Unis 1940, James Algar et al.) se distingue, par exemple, par des niveaux de saturation extrêmement élevés (tracé du milieu), comme c'est d'ailleurs le cas également pour un autre film d'animation, DIE ABENTEUER DES PRINZEN ACHMED [LES AVENTURES DU PRINCE AHMAD] (Allemagne 1925, Lotte Reiniger et Carl Koch) dans le premier tracé à droite. Les titres des films, les segments et les captures d'écran, en combinaison avec une description de la séquence, sont de nouveau affichés par le truchement d'une fonction de survol (figure 27).

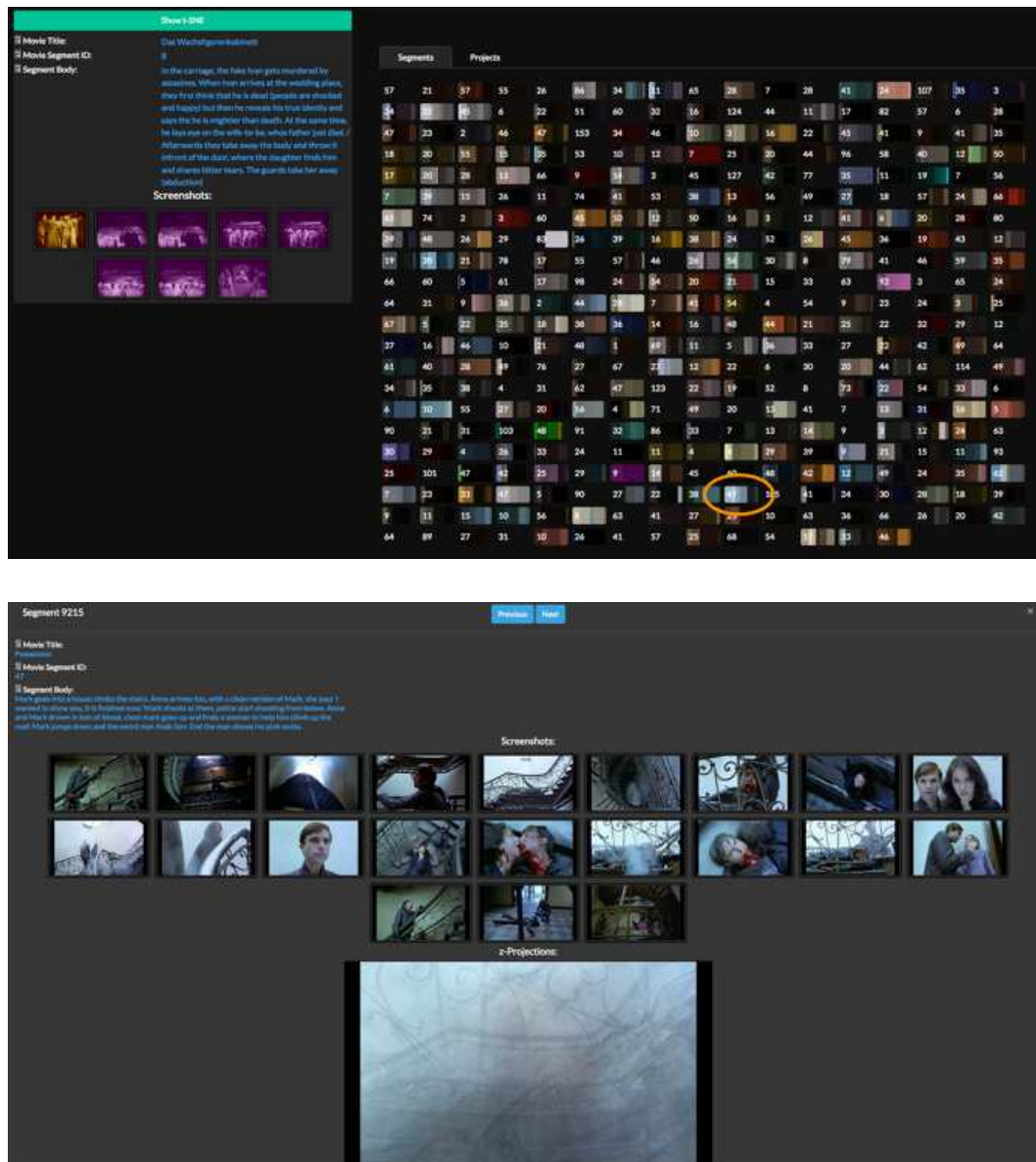


Figure 27. Sur la page de requête de VIAN WebApp, tous les segments et les projets repérés par la requête sont affichés avec une palette de couleurs brutes. En cliquant sur un segment, les captures d'écran, une courte description de la séquence et une visualisation récapitulative s'affichent. POSSESSION (France/Allemagne 1981, Andrzej Żuławski).

8.1 Problèmes

Au cours du développement de ces méthodes de visualisation au niveau du corpus, nous avons rencontré des difficultés pour obtenir des images nettes. L'un des problèmes était lié au flou des concepts qui générant un niveau de bruit assez élevé, comme expliqué dans la Section 7. Or, le problème le plus persistant que nous avons identifié est la prédominance des schémas de couleurs monochromes dans les visualisations LAB, de la même manière que dans les tracés d'images par film discutés dans la Section 7.1. En

raison des niveaux élevés de chromaticité dans certaines captures d'écran monochromes par rapport aux effets de moyenne avec des teintes variées, ces images se démarquent toujours et faussent ainsi le résultat. Cet effet est encore plus fort dans les tracés d'images qui représentent des données au niveau du corpus, à cause des variations dans le traitement de la couleur des différents films.

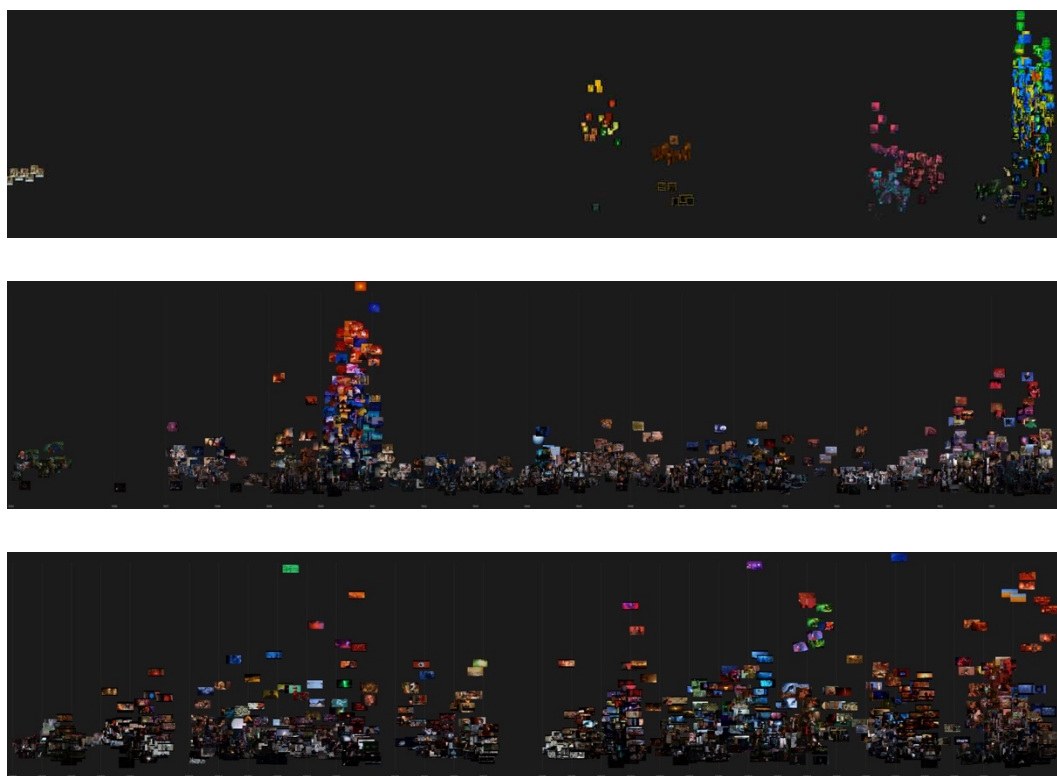


Figure 28. Tracés *Color_dY* de l'évolution de la saturation (axe y) en fonction du temps (axe x) pour les lumières colorées, 1895-1930 (en haut), 1930-1955 (au milieu), 1955-1995 (en bas)

8.2 Leçons tirées

L'une des premières mesures que nous avons prises a été de nettoyer les données. Deuxièmement, nous avons intégré les tracés de points détaillés dans la Section 7, dans des visualisations de corpus. L'un des aspects les plus utiles de la visualisation au niveau du corpus est l'intégration des segments temporels, y compris les descriptions de séquences et les captures d'écran dans une barre latérale à côté des tracés. La relation entre les différents mots-clés devient complexe dans un corpus aussi grand, nous devons donc élaborer davantage de types de visualisations afin de communiquer la corrélation et les liens entre les concepts, les caractéristiques de couleur qui leur sont associées et la distribution

temporelle en fonction du temps en tracés d'images et tracés de points de palette.

9 Variations spatiales, identification et analyse des motifs et des textures

De manière générale, les couleurs sont définies, en termes conceptuels, en fonction de la teinte, de la saturation et de la clarté. Or, du point de vue de la perception, il existe de nombreux autres facteurs qui influencent l'apparence des couleurs et produisent un effet correspondant à notre perception des couleurs (Katz 1911, 1930 ; Hurlbert 2013). L'une des caractéristiques les plus importantes de l'apparence des couleurs, pourtant négligée jusqu'à aujourd'hui, est la *variation spatiale*. Par cela, nous entendons la variation des teintes en fonction de la fréquence spatiale dans une image donnée. Ces variations sont liées à plusieurs facteurs. La complexité de l'image peut être provoquée par des compositions d'image denses, encombrées de nombreux petits détails, aux teintes différentes ou similaires. Des foules énormes de personnages habillés en couleurs différentes constituent un type de sujet qui provoque un degré élevé de variation spatiale. Un autre type comprend des compositions d'image en couches avec des occlusions générées par des objets posés au premier plan.

La complexité visuelle est liée à l'analyse de la texture dans la mesure où les variations spatiales peuvent être une caractéristique qui affecte la lisibilité des compositions d'images (pour une approche en humanités numériques de l'étude de la composition et du style des images, voir Benini et al. 2016). Un facteur supplémentaire est la distribution des teintes, avec un niveau élevé de teintes variables ajoutant à la complexité visuelle. Parallèlement, une distribution de couleurs extrêmement uniforme peut réduire la lisibilité, si elle combinée avec un faible degré de variations spatiales et/ou avec de l'obscurité. La séparation et l'attribution des couleurs sont des indices forts pour la reconnaissance d'objets et pour le repérage de séquences (Hansen et Gegenfurtner 2017).

Dans nos analyses esthétiques, la distinction entre motifs et textures a été fondamentale dès le départ, pour plusieurs raisons. Les *motifs* dénotent des variations de surface sur la base de l'attribution des couleurs – par exemple, des motifs imprimés ou tissés sur des tissus, des surfaces

peintes avec des motifs tels que les papiers peints, etc. En revanche, les *textures* renvoient aux variations de surface tridimensionnelles, telles que les tricots, les rochers, les murs en briques, ainsi que les structures en rondins de bois brutes et non polies. Elles abordent invariablement la perception tactile (Liu, Lughofer / Zeng 2015, Zuo et al. 2016).

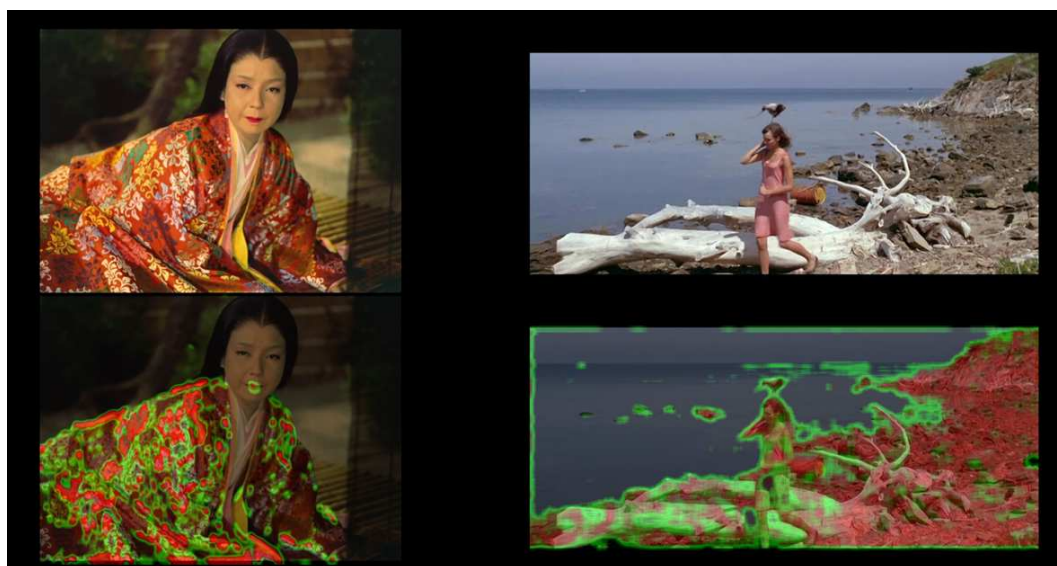


Figure 29. Motifs (à gauche) dans JIGOKUMON (Japon 1953, Teinosuke Kinugasa) vs textures (à droite) dans PIERROT LE FOU (France 1965, Jean-Luc Godard)

L'une des hypothèses directrices de notre recherche a été une connexion forte entre la matérialité des pellicules couleur et les propriétés matérielles de la *diégèse*, l'univers spatiotemporel dépeint dans un film dont les matériaux sont choisis et gérés par le design des costumes et du décor. Par exemple, l'impression demi-teinte telle qu'appliquée dans le traitement par transfert de colorant Technicolor No. III à V manque de définition, car il est difficile d'enregistrer de manière parfaite deux ou trois colorants d'impression, ce qui serait une condition préalable à la résolution spatiale des variations de couleur à petite échelle. Par conséquent, nous nous attendions à ce que ces films omettent des motifs dans leur traitement de la couleur. En revanche, les films teintés manquent de variation spatiale au niveau des teintes car ils sont colorés uniformément en étant immergés dans des bains de teinture – voir *Timeline of Historical Film Colors* (Flueckiger 2012 et suiv.).

Il existe également un lien fort entre, d'une part, les modes affectifs de la perception du film et la complexité visuelle, et, d'autre part, la

lisibilité réduite. Par exemple, la complexité de l'image peut accroître considérablement dans des séquences qui véhiculent un sentiment de stress. La juxtaposition en plusieurs couches d'images, considérée comme la pierre angulaire du système hollywoodien, constitue l'un des problèmes majeurs pour la détection automatique des films. Cependant, comme l'a montré la recherche de l'une des membres de notre équipe, Michelle Beutler, le système hollywoodien lui-même était beaucoup moins normatif qu'on ne le supposait auparavant. L'augmentation des propriétés tactiles et des subjectivités affectivement chargées dans les films à partir des années 1960 est au centre de l'étude de Bregt Lameris sur les couleurs et les affects au cinéma (voir Lameris 2019). Dans l'étude de Joëlle Kost sur les types de pellicule chromogène, la complexité visuelle est l'un des principaux sujets, étant donné qu'elle est liée à l'amélioration de la résolution de ces pellicules.

En formant un réseau d'apprentissage profond pour réaliser une segmentation de sous-figures par pixel à l'aide de l'ensemble de données LIP (Gong et al. 2017), nous serons en mesure d'analyser les deux fonctionnalités¹⁸.

Une possibilité pour l'évaluation de la fréquence spatiale au sein d'un plan est d'utiliser un algorithme de détection de contours – l'idée étant que plus il y a d'arêtes, plus la région est dense. Cela s'est déjà avéré être une mesure solide pour la complexité spatiale, bien qu'elle ne couvre pas seulement la variation liée à la teinte et à la chrominance. VIAN visualise actuellement trois mesures différentes sous forme de carte thermique sur le lecteur : la densité des bords imbriqués, la luminance par pixel et la variance d'un canal $a*b^*$.

9.1 Problèmes

Distinguer entre les motifs et la texture n'est pas une tâche computationnelle triviale. Une approche naïve consisterait à assumer que la variance de la luminance indique généralement une qualité tactile,

¹⁸ Gaudenz Halter avait développé des outils d'apprentissage profond pour l'analyse de textures, modèles et matériels. Son travail est documenté dans son mémoire de master *Perceptual Texture Features in Moving Image. An Exploratory Visualization Approach* (2020).

tandis qu'une variance élevée de la teinte et de la chrominance indiquerait des motifs. Néanmoins, puisque les motifs ne sont pas exclus de la variance élevée dans le canal de la luminance, cette approche ne donne pas de résultats précis. En outre, il arrive souvent que certains matériaux possédant une qualité tactile pour les êtres humains ne soient pas significativement différents des surfaces plates, en termes numériques. La covariance de la fréquence spatiale, des valeurs de couleur (teinte, clarté, saturation) et des textures par rapport aux motifs est étroitement liée aux processus d'ordre supérieur de la perception visuelle humaine, tels que la mémoire des couleurs et l'intégration multimodale, c'est-à-dire le lien entre l'expérience tactile et la perception visuelle et auditive. Comme indiqué précédemment dans l'étude de Flueckiger sur le traitement du son, les propriétés matérielles sont souvent mieux détectées via leurs signaux acoustiques. Il serait utile d'inclure le son dans une itération future et plus élaborée du système.

9.2 Leçons tirées

La fréquence spatiale ainsi que la différenciation et l'évaluation des motifs et des textures font toujours partie de nos recherches actuelles. La complexité visuelle est l'un des facteurs les plus importants en matière de style et d'évolutions diachroniques. Ainsi, nous avons intégré dans le projet une étude de suivi oculaire, afin d'obtenir des informations empiriques sur le sujet (voir Smith / Mital 2013 ; Rubo / Gamer 2018). L'étude a été menée par Miriam Loertscher en coopération avec Bregt Lameris. Pour cette étude, nous avons choisi un ensemble de séquences représentatives pour différents types de composition d'image et de complexité – par exemple, le type net sans motifs et textures comme dans *UNE FEMME EST UNE FEMME*, le type du « monde saturé d'objets » comme dans *MORTE A VENEZIA* [MORT À VENISE] avec des compositions d'image en plusieurs couches excessivement denses, ou *SAYAT NOVA* [LA COULEUR DE LA GRENADE], un film qui fonctionne avec de nombreuses textures et variations matérielles, en excluant souvent la présence humaine.

Les résultats sont actuellement en cours de traitement ; or, à partir d'un bref aperçu des cartes thermiques, les parties d'image avec des variations à petite échelle détournent le plus le regard des spectateurs des

personnages et, surtout, des visages. Comme le déclarent Rubo et Gamer (2018 : 1) : « L'influence des stimuli sociaux et de la saillance visuelle de bas niveau sur les mouvements oculaires n'a été étudiée que récemment au sein du même ensemble de données, et rarement en juxtaposition directe. Au cours de la perception faciale, il a été montré que le diagnostic des régions du visage pour les expressions d'émotion a reçu une attention accrue indépendamment de sa saillance physique de bas niveau ».



Figure 10. Cartes thermiques de l'étude de suivi oculaire : une forte séparation figure-fond et une absence de variation spatiale dans l'arrière-plan dans *UNE FEMME EST UNE FEMME* (en haut et à gauche) ; des textures et des motifs en combinaison avec l'absence des êtres humains dans *SAYAT NOVA* (en haut et à droite) ; et le monde saturé d'objets avec un niveau élevé d'obstruction et de complexité visuelle dans *MORTE A VENEZIA* (en bas et au centre). Cette étude a été menée par Miriam Loertscher et Bregt Lameris, ERC Advanced Grant *FilmColors*.

Les hypothèses qui en résultent doivent être testées pour identifier des régions de fréquence spatiale élevée et comparées aux données récoltées manuellement afin d'évaluer si ces régions représentent des motifs ou des textures.

10 Conclusion

Dans le présent article, nous avons discuté du potentiel et des limites des outils numériques pour l'analyse de l'esthétique et du récit filmiques basés sur le cas de la recherche sur la technologie et l'esthétique des couleurs des films. En suivant l'argument central établi dans l'introduction – à savoir que de tels outils nécessitent une base théorique solide, l'interprétation humaine, une discussion continue et une réflexion approfondie sur les préconceptions intégrées dans les outils eux-mêmes –, nous avons exploré diverses approches pour relier la perspective des sciences humaines aux méthodes de la science des données et de l'informatique.

Nos recherches ont démontré le besoin d'employer un large éventail de méthodes fines d'analyse et de visualisation afin d'éviter les pièges de généralisations infondées et d'études anecdotiques. Sur le plan négatif d'un si grand ensemble de données et d'une gamme sophistiquée de concepts théoriques et analytiques, il existe une complexité et un bruit considérables qui tendent à obscurcir les résultats. Nous avons trouvé que pour chacune des questions de recherche, nous devons prendre en considération la gamme complète des visualisations et réévaluer les résultats au cas par cas.

Comparés aux approches traditionnelles, principalement discursives, de l'esthétique, de la technologie et de la narratologie des couleurs des films, les outils des humanités numériques convertissent les preuves créées par la cartographie des résultats en un tableau instantanément accessible de représentations visuelles. En associant l'annotation et l'interprétation humaines détaillées à ces représentations visuelles, le flux opérationnel intégré, composé du logiciel d'analyse visuelle VIAN en combinaison avec le portail de *crowdsourcing* VIAN WebApp, a formé un écosystème complet pour l'étude de l'esthétique et de la narration filmiques. Ainsi, il élargit considérablement les méthodes établies dans les études cinématographiques.

Dans les projets de coopération actuels ou planifiés, nous cherchons à explorer et à élargir cette approche au-delà de la problématique des couleurs des films, vers l'étude des méthodes d'enseignement, de recherche et des sciences participatives.

11 Remerciements

Ce projet est financé par l'ERC (*European Research Council*) dans le cadre du programme de recherche et d'innovation *Horizon 2020* de l'Union européenne (accord de subvention No 670446 *FilmColors*). Les analyses ont été réalisées par la chercheuse principale (CP) Barbara Flueckiger, la chercheuse postdoctorale Bregt Lameris, les doctorantes Olivia Kristina Stutz, Michelle Beutler et Joëlle Kost, et trois étudiant.e.s assistant.e.s, Manuel Joller, Valentina Romero et Ursina Früh. Le développement de VIAN et de VIAN WebApp a été soutenu par le Laboratoire de Visualisation et Multimédia VMML de l'Université de Zurich, dirigé par Renato Pajarola avec Enrique Paredes et Rafael Ballester-Ripoll.

References

- ACLAND, C.R, HOYT, E., *The Arclight Guidebook to Media History and the Digital Humanities*, Sussex UK, Reframe, 2016,
<http://reframe.sussex.ac.uk/reframebooks/archive2016/the-arclight-guidebook-to-media-history-and-the-digital-humanities/>
- ARNOLD, T., TILTON, L., « Distant Viewing Toolkit. A Python Package for the Analysis of Visual Culture », *Journal of Open Source Software*, 45,5, 2020, pp. 1–6, <https://doi.org/10.21105/joss.01800>
- ARNOLD, T., TILTON, L. « Enriching Historic Photography with Structured Data using Image Region Segmentation », *Proceedings of the First Artificial Intelligence for Historical Image Enrichment and Access (AI4HI)*, Marseille, France, Association for Computational Linguistics, 2020b.
- BEAU, F. « La solitude du technobole. Puissance politique des effets spéciaux », *CinémAction*, n° 102, 2001, dossier « Du trucage aux effets spéciaux », pp. 196–207.
- BENINI, S., SVANERA, M., ADAMI, N., LEONARDI, R. and KOVÁCS, A. B. « Shot Scale Distribution in Art Films », *Multimedia Tools Appl.*, n° 75, 2016, pp. 16499–16527, <https://doi.org/10.1007/s11042-016-3339-9>

- BOHN, C. « Clothing as Medium of Communication », *Soziale Systeme*, 6,1, 2000, pp. 111–137.
- BORDWELL, D. « Historical Poetics of Cinema », in : PALMER, B., (ed.), *The Cinematic Text. Methods and Approaches*, New York, ASM Press, 1989, pp. 370–398.
- BORDWELL, D., STAIGER, J. and THOMPSON, K., *The Classical Hollywood Cinema. Film Style and Mode of Production to 1960*, London, Routledge, 1985.
- BRODBECK, F., *Cinemetrics. Film Data Visualization*, 2011, <http://cinemetrics.fredericbrodbeck.de/>
- BURGHARDT, M., KAO, M. and WOLFF, C. « Beyond Shot Lengths. Using Language Data and Color Information as Additional Parameters for Quantitative Movie Analysis », *Digital Humanities 2016. Conference Abstracts*, Kraków, Jagiellonian University & Pedagogical University, 2016, pp. 753–755.
- BURGHARDT, M., HAFNER, K., EDEL, L., KENAAN, S.-L. and WOLFF, C., « An Information System for the Analysis of Color Distributions in MovieBarcodes », in : GÄDE, M.(ed.), *Proceedings of the 15th Int. Symp. of Information Science (ISI 2017)*, Berlin, Germany, 13th–15th March 2017, Glückstadt, Verlag Werner Hülsbusch, 2017, pp. 356–358., <https://epub.uni-regensburg.de/35682/>
- CASEY, M., WILLIAMS, M. « Action. Audio-visual Cinematics Toolbox for Interaction, Organization, and Navigation of Film » (White Paper Report No. HD5139411), Hanover, Dartmouth College, 2014, <https://hcommons.org/deposits/item/hc:12153/>
- CUTTING, J.E., « Perception, Attention, and the Structure of Hollywood Film », 2016, <http://people.psych.cornell.edu/~jec7/currerearch.htm>
- CUTTING J.E., BRUNICK, K. L. et CANDAN, A. « Perceiving Event Dynamics and Parsing Hollywood Films », *Journal of Experimental Psychology, Advance online publication*, 2012, <https://doi.org/10.1037/a0027737>
- DYER, R., *Pastiche*, New York, Routledge, 2006.
- ELAN. <https://tla.mpi.nl/tools/tla-tools/elan/>

- FERGUSON, K.L. « Western Roundup », 2013,
<http://typecast.qwriting.qc.cuny.edu/2013/10/07/western-roundup/>
- FERGUSON, K.L. « Volumetric Cinema », 2015, <https://vimeo.com/119790662>
- FERGUSON, K.L. « The Slices of Cinema. Digital Surrealism as Research Strategy », in: ACLAND, C. R. et HOYT, E. (Éds.), *The Arclight Guidebook to Media History and the Digital Humanities*, Reframe Books, 2016, pp. 270–299, <http://projectarclight.org/book/>
- FLUECKIGER, B., « Die Vermessung ästhetischer Erscheinungen », *Zeitschrift für Medienwissenschaft*, n° 5, 2011, pp. 44–60.
- FLUECKIGER, B. *Timeline of Historical Film Colors*, 2012-2020,
<https://filmcolors.org/>
- FLUECKIGER, B., « A Digital Humanities Approach to Film Colors », *The Moving Image*, 17.2, 2017, pp. 71–94,
<https://www.jstor.org/stable/10.5749/movingimage.17.2.00710>
- FLUECKIGER, B., EVIRGEN, N., PAREDES, E. G., BALLESTER-RIPOLL, R. et PAJAROLA, R. « Deep Learning Tools for Foreground-Aware Analysis of Film Colors », 2017, <https://avindhsig.wordpress.com/deep-learning-tools-for-foreground-aware-analysis-of-film-colors/>
- FLUECKIGER, B., HALTER, G., « Building a Crowdsourcing Platform for the Analysis of Film Colors », *The Moving Image*, 18.1, 2018, pp. 80–83.
<https://doi.org/10.5749/movingimage.18.1.0080>
- FLUECKIGER, B., « The Material Aesthetics of Len Lye's Experimental Color Films in the 1930s », conférence présentée au *Len Lye Symposium*, Tinguely Museum Bâle, 2019.
- FLUECKIGER, B., « Film Colors. Materiality, Technology, Aesthetics », in: FLUECKIGER, B., HIELSCHER, E. et WIETLISBACH, N. (Éds.), *Color Mania. The Material of Color in Photography and Film*, Zurich ; Winterthur, Lars Müller ; Fotomuseum Winterthur, 2020, pp. 17–49.
- FOSSATI, G., « When the Cinema Was Colored », in : BERRIATÚA, L., BRUSATIN, M., DESMET, N., FORNAROLI, E., FOSATTI, G., LAVEDRINE, B., et al., *Tutti i colori del mondo. Il colore nei mass media tra 1900 e 1930*, Reggio Emilia, Diabasis, 1998, pp. 121–132.
- GENETTE, G. *Figures III*, Paris, Seuil, 1972.

GENETTE, G., *Nouveau Discours du récit* [1ère édition], Paris, Le Seuil, 1983.

GIUNTI, L. *Problemi dell'analisi del testo di finzione audiovisivo. Verifica e sviluppo di un modello analitico e interpretativo con strumenti digitali*, Università degli Studi di Pisa, 2010,
https://etd.adm.unipi.it/theses/available/etd-10172012-200229/unrestricted/TESI_caricamento.pdf

GIUNTI, L. « L'analyse du film à l'ère numérique. Annotation, geste analytique et lecture active », *Cinéma & Cie*, 14, 22/23, 2014, pp. 127–143,
http://www.academia.edu/13710402/L_analyse_du_film_a_l_ere_numerique._Annotation_geste_analytique_et_lecture_active

GONG, K., LIANG, X., ZHANG, D., SHEN, X. et LIN, L., « Look into Person. Self-Supervised Structure-Sensitive Learning and a New Benchmark for Human Parsing », 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6757–6765, 2017,
<https://doi.org/10.1109/CVPR.2017.7150>

GRUBER, K., WURM, B., KROPF, V., (Éds.), *Digital Formalism. Die kalkulierten Bilder des Dziga Vertov*, Vol. 55, 3, Wien, Böhlau Verlag, 2009.

HAGBERG, A. A., SCHULT, D. A., SWART, P. J., « Exploring Network Structure, Dynamics, and Function using NetworkX », *Proceedings of the 7th Python in Science Conference*, pp. 11–15, Pasadena, CA USA, 2008.

HAHN, S., « Filmprotokoll Revisited. Ground Truth in Digital Formalism », in : GRUBER, K. WURM, B. et KROPF, V. (Éds.), *Digital Formalism: Die kalkulierten Bilder des Dziga Vertov*, Wien, Böhlau Verlag, 2009, pp. 129–136.

HALTER, G., BALLESTER-RIPOLL, R., FLUECKIGER, B. et PAJAROLA, R., « VIAN. A Visual Annotation Tool for Film Analysis », *Computer Graphics Forum*, 38.3, 2019, pp. 119–129, <https://doi.org/10.1111/cgf.13676>

HALTER, G., *Perceptual Texture Features in Moving Image. An Exploratory Visualization Approach*. Zurich: University of Zurich, 2020.

HANSEN, T., GEGENFURTNER, K. R., « Color Contributes to Object-contour Perception in Natural Scenes », *Journal of Vision*, 17.3, 2017, pp. 14–14,
<https://doi.org/10.1167/17.3.14>

HEFTBERGER, A. « Die Verschmelzung von Wissenschaft und Filmchronik. Das Potenzial der Reduktionslosen Visualisierung am Beispiel von Das Elfte Jahr und Der Mann mit der Kamera von Dziga Vertov », *La Visualisation des Données en Histoire = Visualisierung von Daten in der*

- Geschichtswissenschaft*, 2015, pp. 229–263.
- HEFTBERGER, A., *Kollision der Kader. Dziga Vertovs Filme, die Visualisierung ihrer Strukturen und die Digital Humanities*, München, Edition Text + Kritik, 2016.
- HEFTBERGER, A., *Digital Humanities and Film Studies. Visualising Dziga Vertov's Work*, Berlin, Springer International Publishing, 2018.
- HURLBERT, A., « The Perceptual Quality of Color », in : ALBERTAZZI, L., (éd.), *Handbook of Experimental Phenomenology. Visual Perception of Shape, Space and Appearance*. New York, John Wiley & Sons, Ltd SN, 2013, pp. 369–394, <http://dx.doi.org/10.1002/9781118329016.ch15>
- ITTEN, J., *Kunst der Farbe. Studienausgabe*, Ravensburg, Ravensburger Buchverlag, 1970.
- JAMESON, F., *Postmodernism. Or, the Cultural Logic of Late Capitalism*, London, Verso, 1991.
- KATZ, D., *Die Erscheinungsweisen der Farben und ihre Beeinflussung durch die individuelle Erfahrung*, London, Kegan Paul, Trench, Trubner & Co. Ltd., 1911.
- KATZ, D., *Der Aufbau der Farbwelt*, Leipzig, Johann Ambrosius Barth, 1930.
- LAMERIS, B., « Hallucinating Colours. Psychedelic Film, Technology, Aesthetics and Affect », *Cinéma & Cie*, Special Issue « Cinema and Mid-Century Colour Culture », 32,19, 2019, pp. 85-97.
- LEDIG, E., ULLMANN, G., « Rot wie Feuer, Leidenschaft, Genie, Wahnsinn. Zu einigen Aspekten der Farbe im Stummfilm », in : LEDIG, E. (éd.), *Der Stummfilm. Konstruktion und Rekonstruktion*, München, Schaudig, Bauer, Ledig, 1988, pp. 89–116.
- LIU, J., LUGHOFFER, E. et ZENG, X., « Aesthetic Perception of Visual Textures. A Holistic Exploration Using Texture Analysis, Psychological Experiment, and Perception Modeling », *Frontiers in Computational Neuroscience*, 9, 2015, pp. 1–14, <https://doi.org/10.3389/fncom.2015.00134>
- LONG, J., SHELHAMER, E., et DARRELL, T., « Fully Convolutional Networks for Semantic Segmentation », pp. 3431–3440, conférence présentée aux *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, <https://arxiv.org/abs/1411.4038>

- MANOVICH, L., « How to Compare One Million Images? », in : BERRY, D., (éd.), *Understanding Digital Humanities*, London, Palgrave Macmillan UK, 2012, pp. 249–278.
- MANOVICH, L., « Kino-Eye in Reverse. Visualizing Cinema », in : GEIGER, J. et LITTAU, K. (éds.), *Cinematicity in Media History*, Edinburgh, Edinburgh University Press, 2013, pp. 211–234.
- MANOVICH, L., « Visualizing Vertov », *Russian Journal of Communication*, 5,1, 2013, pp. 44–55, <https://doi.org/10.1080/19409419.2013.775546>
- MELGAR ESTRADA, L., HIELSCHER, E., KOOLEN, M., OLESEN, C. G., NOORDEGRAAF, J. et BLOM, J., « Film Analysis as Annotation. Exploring Current Tools », *The Moving Image: The Journal of the Association of Moving Image Archivists*, 17.2, 2017, pp. 40–70, <https://doi.org/10.5749/movingimage.17.2.0040>
- OLESEN, C. G., *Film History in the Making. Film Historiography, Digitised Archives and Digital Research Dispositifs* (thèse de doctorat). University of Amsterdam, Amsterdam, 2017, <https://dare.uva.nl/search?identifier=ad68a275-e968-4fce-b91e-4783cd69686c>
- OLESEN, C. G., MASSON, E., GORP, J. van, FOSSATI, G. et NOORDEGRAAF, J., « Data-Driven Research for Film History. Exploring the Jean Desmet Collection », *The Moving Image*, 16,1, 2016, pp. 82–105, <https://muse.jhu.edu/article/640569>
- OLESEN, C. G., GORP, J. v., FOSSATI, G., « Datasets and Colour Visualizations for Data-Driven Film History. A Demonstrator of EYE's Jean Desmet Collection », 2016, <http://www.create.humanities.uva.nl/results/desmetdatasets/>
- PLANTINGA, C., *Moving Viewers. American Film and the Spectator's Experience*, Berkeley, University of California Press, 2009.
- REDMON, J., DIVVALA, S., GIRSHICK, R. et FARHADI, A., « You Only Look Once. Unified, Real-Time Object Detection », arXiv:1506.02640 [cs], 2015, <http://arxiv.org/abs/1506.02640>
- REYES-GARCÍA, E., « On Visual Features and Artistic Digital Images », conférence présentée à *VRC'13 Proceedings of the Virtual Reality International Conference*, Laval Virtual, Laval Franc New York, ACM, 2013, <http://dl.acm.org/citation.cfm?id=2466835>

- REYES-GARCÍA, E., « Explorations in Media Visualization », conférence présentée à *Extended Proceedings of the 25th ACM Conference on Hypertext and Hypermedia Hypertext*, New York, ACM, 2014, http://ceur-ws.org/Vol-1210/datawiz2014_11.pdf
- REYES-GARCÍA, E., *The Image-interface. Graphical Supports for Visual Information*, Hoboken, NJ, Wiley-ISTE, 2017.
- REYES-GARCÍA, E., BOUHAI, N. (Éds.), *Designing Interactive Hypermedia Systems* (première édition), Wiley-ISTE, 2017.
- ROSEBROCK, A., « How-To. OpenCV and Python K-Means Color Clustering », 2014, <http://www.pyimagesearch.com/2014/05/26/opencv-python-k-means-color-clustering/>
- ROSS, J., « Colour Analysis Tools in ImageJ », 2007, https://www.fmhs.auckland.ac.nz/assets/fmhs/sms/biru/docs/Colour_Analysis_Tools_in_ImageJ.pdf
- ROTHER, C., KOLMOGOROV, V. et BLAKE, A., « GrabCut. Interactive Foreground Extraction using Iterated Graph Cuts », *ACM Transactions on Graphics (SIGGRAPH)*, 2004.
- RUBO, M., GAMER, M. « Social Content and Emotional Valence Modulate Gaze Fixations in Dynamic Scenes », *Scientific Reports*, 8.1, 2018, pp. 1–11. <https://doi.org/10.1038/s41598-018-22127-w>
- SMITH, M., *Engaging Characters. Fiction, Emotion, and the Cinema*, Oxford ; New York, Oxford University Press, 1995.
- SMITH, M., MITAL, P. K., « Attentional Synchrony and the Influence of Viewing Task on Gaze Behavior in Static and Dynamic Scenes », *Journal of Vision*, 13.8, 2013, pp. 1–24, <https://doi.org/10.1167/13.8.16>
- STUTZ, O. K., *Algorithmische Farbfilmästhetik. Historische sowie experimentell-digitale Notations- und Visualisierungssysteme des Farbfilms im Zeichen der Digital Humanities 2.0 und 3.0* (mémoire de master), Universität Zürich, Zürich, 2016, http://www.film.uzh.ch/dam/jcr:bed543b6-4a67-4ff8-8f51-b85a739417d5/MA_AlgorithmischeFarbfilmaesthetik_Stutz_HS2016_def.pdf
- THOMPSON, K., « The Concept of Cinematic Excess », in : ROSEN, P., BRAUDY, L. et COHEN, M. (Éds.), *Film Theory and Criticism. Introductory Readings*, New York, Columbia University Press, 1986, pp. 487–498.

THOMPSON, K., *Breaking the Glass Armor*, New Jersey, Princeton University Press, 1988.

VÄNSKÄ, A., « Gender and Sexuality », in : PALMER, A. (éd.), *A Cultural History of Dress and Fashion*, Bloomsbury Academic, 2017, pp. 107–129.

WILKENING, A., « Die Restaurierung von Das Cabinet des Dr. Caligari », *VDR Beiträge zur Erhaltung von Kunst- und Kulturgut*, 2, 2014, pp. 27–47.

WILLIAMS, M., « Networking Moving Image History. Archives, Scholars, and the Media Ecology Project », in : HOYT, E. et ACLAND, C. R. (Éds.), *The Arclight Guidebook to Media History and the Digital Humanities*, Sussex UK, ReFrame, 2016, pp. 334–345,
<http://reframe.sussex.ac.uk/reframebooks/archive2016/the-arclight-guidebook-to-media-history-and-the-digital-humanities/0>

ZHAO, H., SHU, J., QI, X., WANG, X. et JIA, J., « Pyramid Scene Parsing Network », arXiv:1612.01105 [cs], 2016, <http://arxiv.org/abs/1612.01105>

ZUO, H., JONES, M., HOPE, T., et JONES, R., « Sensory Perception of Material Texture in Consumer Products », *The Design Journal*, 19.03, 2016, pp. 405–427, <https://doi.org/10.1080/14606925.2016.11493180>

À propos des auteur.e.s

Barbara Flueckiger

baflueckiger@gmail.com

ERC Advanced Grant FilmColors, University of Zurich

Barbara Flueckiger est professeure en études de cinéma à l'Université de Zurich depuis 2007. Avant ses études en théorie et histoire du cinéma, elle a travaillé à l'international en tant que professionnelle du cinéma. Elle a publié deux ouvrages : « Sound Design » et « Visual Effects ». Depuis 2001, elle a conçu et dirigé plusieurs projets de recherche sur l'esthétique et les technologies du film. Ses travaux récents portent sur la numérisation et l'archivage de films, menés en collaboration avec l'industrie du cinéma et des archives. En 2015 elle a été reconnue avec le prestigieux Advanced Grant de l'European Research Council (ERC) pour un projet de recherche qui s'intéresse aux relations entre les technologies et l'esthétique de couleurs des films.

Gaudenz Halter

University of Zurich

Gaudenz Halter est doctorant en informatique à l'Université de Zurich. Pendant ses études de master en data science, il a développé des applications d'analyse et d'annotation de films dans le cadre du projet ERC Advanced Grant *FilmColors*.

Faire image

Inscriptions, espaces et formes numériques

À l'ère du numérique, l'image est le support prépondérant grâce auquel nous construisons nos interactions sociales, culturelles, économiques et politiques. Les approches émergentes de l'étude et de l'analyse des images nous semblent donc appeler un regard neuf, basé sur des croisements interdisciplinaires. Afin de tisser des liens entre différentes disciplines et d'interroger ce qui fait image à l'ère du numérique, nous avons organisé un symposium international « Faire image : inscriptions, espaces et formes numériques » le 28 février 2020 à la Maison de la recherche de l'Université Paris 8 Vincennes-Saint-Denis. Ce volume réunit des contributions issues du colloque international, et propose une version révisée et approfondie des communications. Il s'enrichit également de l'apport de nos débats passionnés, tous les textes ayant été rédigés après l'événement.

SOUS LA DIRECTION DE
Gwen Le Cor et
Everardo Reyes

Octaviana

Bibliothèque numérique de l'Université Paris 8

93526 Saint-Denis cedex

<https://octaviana.fr/>

 octaviana

